

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 22.4894 Seconds
(without alignments)
64.134 Million cell updates/sec

Title: US-09-856-086-4
Perfect score: 36
Sequence: 1 ELEDKON 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertibrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	386	4	Q8TCR7
2	32	88.9	357	10	Q9SVY8
3	32	88.9	403	5	Q9U250
4	32	88.9	1170	16	Q9XOR4
5	32	88.9	1418	5	Q01837
6	32	88.9	1420	5	Q960A2
7	31	86.1	180	5	Q9VZ73
8	31	86.1	366	2	Q9RH82
9	31	86.1	499	9	Q9A2C6
10	31	86.1	657	10	Q93Y14
11	31	86.1	674	10	Q23544
12	30	83.3	286	17	Q9UWV0
13	30	83.3	293	2	Q68333
14	30	83.3	293	2	Q93K78
15	30	83.3	348	17	Q8TNU6
16	30	83.3	370	2	Q9Z9V6

17	30	83.3	399	10	Q9LJ29	Q9LJ29 arabidopsis
18	30	83.3	400	16	Q9KGA7	Q9KGA7 bacillus ha
19	30	83.3	547	16	Q9A1N2	Q9A1N2 streptococ
20	30	83.3	564	2	Q9X2Y5	Q9X2Y5 bacillus an
21	30	83.3	599	2	Q8RSR4	Q8RSR4 lactobacill
22	30	83.3	715	16	P73259	P73259 synecocyst
23	30	83.3	978	16	Q9PNK9	Q9PNK9 campylobact
24	30	83.3	1146	5	Q9N8Q9	Q9N8Q9 trypsinosoma
25	30	83.3	1228	5	O17566	O17566 caenorhabd1
26	30	83.3	1410	5	Q23371	Q23371 caenorhabd1
27	30	83.3	1713	5	Q9VRG4	Q9VRG4 drosophila
28	30	83.3	2701	6	Q8WN96	Q8WN96 bos taurus
29	29	80.6	146	12	Q65210	Q65210 african sw1
30	29	80.6	189	16	Q53647	Q53647 staphylococ
31	29	80.6	263	16	O50958	O50958 borrelia bu
32	29	80.6	272	12	Q93GV4	Q93GV4 rice grassy
33	29	80.6	281	16	Q9KAU6	Q9KAU6 bacillus ha
34	29	80.6	315	17	Q978U8	Q978U8 thermoplasma
35	29	80.6	346	5	Q9VE86	Q9VE86 drosophila
36	29	80.6	353	5	Q961E1	Q961E1 drosophila
37	29	80.6	358	5	Q22596	Q22596 caenorhabd1
38	29	80.6	402	1	Q9HH69	Q9HH69 methanobact
39	29	80.6	441	10	Q9PLM5	Q9PLM5 arabidopsis
40	29	80.6	457	10	Q94CD1	Q94CD1 arabidopsis
41	29	80.6	486	16	Q9PN48	Q9PN48 campylobact
42	29	80.6	522	5	Q8T202	Q8T202 dictyostell
43	29	80.6	523	5	O18378	O18378 caenorhabd1
44	29	80.6	534	16	Q9K717	Q9K717 vibrio chol
45	29	80.6	569	3	Q9HGQ5	Q9HGQ5 saccharomyc

ALIGNMENTS

RESULT 1	Q8TCR7	PRELIMINARY;	PRT;	386 AA.
ID	Q8TCR7			
AC	Q8TCR7			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Hypothetical 43.8 kDa protein (Fragment).			
GN	DKFZP761K0922.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=AMYGDALA;			
RA	Wambutt R., Heubner D., Meves H.W., Well B., Wiemann S.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AL713644; CAD28456.1; -.			
KW	Hypothetical protein.			
FT	NON_TER			
SO	SEQUENCE	386 AA;	43779 MW;	F293388B200CTB65 CRC64;
Query Match				
Best Local Similarity		100.0%;	Score 36;	DB 4;
Matches	7;	Conservative	0;	Mismatches
			0;	Indels
			0;	Gaps
			0;	
QY	1 ELEDKON 7			
DB	178 ELEDKON 184			
RESULT 2	Q9SVY8	PRELIMINARY;	PRT;	357 AA.
ID	Q9SVY8			
AC	Q9SVY8			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			

DE Hypothetical 41.5 kDa protein.
GN F158870.
OS Arabidopsis thaliana (Mouse-ear cress). Embryophyta; Tracheophyta;
OC Eukaryota; Viridiplantae; Streptophyta; Eudicotyledons; core eudicots; Rosidae;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eudicots II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Benes V., Rechmann S., Borkova D., Ansoerge W., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C., Quettler F., Salanoubat M.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049660; CAB41181.1; -
DR InterPro: IPR000719; Euk_Pkinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Euk_Pkinase; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Hypothetical protein; Transferrase.
SQ SEQUENCE 357 AA; 41497 MW; A/CFS01AB60DF620 CRC64;
Query Match 88.9%; Score 32; DB 10; Length 357;
Best Local Similarity 85.7%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 ELEDKON 7
Db 273 ELEDKON 279
RESULT 3
O90250 PRELIMINARY; PRT; 403 AA.
AC O90250;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Y53H1B.1 protein.
GN Y53H1B.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA White S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; AL132851; CAB60411.1; -
DR InterPro: IPR001507; Endoglilin/CD105.
DR SMART; SM00241; ZP; 1.
SQ SEQUENCE 403 AA; 45346 MW; EECES55419CA8428 CRC64;
Query Match 88.9%; Score 32; DB 5; Length 403;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 ELEDKON 7
Db 329 ELEDKON 335
RESULT 4
O9X0R4

ID O9X0R4 PRELIMINARY; PRT; 1170 AA.
AC O9X0R4;
DT 01-NOV-1999 (TRENBLREL. 12, Created)
DT 01-NOV-1999 (TRENBLREL. 12, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Chromosome segregation SMC protein, putative.
GN TM1182.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MSB / DSM 3109;
RX MEDLINE-99287316; PubMed-10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher R.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
DR EMBL; AE001774; AAD36257.1; -
DR TIGR; TM1182;
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR005289; GTP-binding_dom.
DR InterPro: IPR003405; SMC_C.
DR InterPro: IPR003395; SMC_N.
DR InterPro: IPR002017; Spectrin.
DR Pfam; PF02483; SMC_C; 1.
DR Pfam; PF02483; SMC_N; 1.
DR TIGRfams; TIGR00650; MG442; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 1170 AA; 137587 MW; 81B203B54B336DE4 CRC64;
Query Match 88.9%; Score 32; DB 16; Length 1170;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 ELEDKON 7
Db 422 ELEDKON 428
RESULT 5
O01837 PRELIMINARY; PRT; 1418 AA.
AC O01837;
DT 01-JUL-1997 (TRENBLREL. 04, Created)
DT 01-JUL-1997 (TRENBLREL. 04, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Hypothetical 163.4 kDa protein.
GN B0414.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Sammons L., Wohlmann P., Rohlfing T.;
RT "The sequence of C. elegans cosmid B0414."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.

[3]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL: AF003145; AAB57718.1; -
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Hypothetical protein; Serine/threonine-protein kinase;
 KM Transferrase.
 SQ SEQUENCE 1418 AA; 163392 MW; 20967687701221D5 CRC64;

Query Match 88.9%; Score 32; DB 5; Length 1418;
 Best Local Similarity 85.7%; Pred. No. 3.5e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 Db 1082 ELEDKRN 1088

RESULT 6
 O960A2 PRELIMINARY; PRT; 1420 AA.
 ID O960A2;
 AC 01-DEC-2001 (TRENBLREL. 19, Created)
 DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
 DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
 DE Hypothetical protein B0414.7b.
 GN B0414.7
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OC NCBITaxID=6239;

[1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium."
 RL Science 282:2012-2018(1998).
 CC [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Sammons L., Wohlmann P., Rohlfing T.;
 RT "The sequence of C. elegans cosmid B0414."
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 CC [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "Direct Submission."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF003145; AAK66678.1; -
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN.1.
 KW ATP-binding; Transferrase.
 SQ SEQUENCE 1420 AA; 163667 MW; 6B6BFE0902F021C0 CRC64;

Query Match 88.9%; Score 32; DB 5; Length 1420;
 Best Local Similarity 85.7%; Pred. No. 3.5e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ELEDKON 7
 Db 1082 ELEDKRN 1088

RESULT 7
 O9VZ73 PRELIMINARY; PRT; 180 AA.
 ID O9VZ73;
 AC 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
 DE CG17030 protein (AT26201p).
 GN CG17030.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBITaxID=7227;

[1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durlin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fester C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchem K.A.,
 RA Kimmel B.E., Kodra C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mactel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Moadary C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo K., Pittman G.S., Pan S., Pollard J., Puig V., Reese M.G.,
 RA Relier K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 CC [2]
 RN SEQUENCE FROM N.A.
 RP Stapleton M., Brocksstein P., Hong L., Agbayani A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacle J., Pargass V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celinker S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER

CC PROTEINS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +
 CC DIPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
 CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
 CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL; AE003483; AAF47953.1; -
 DR EMBL; AY089444; AAL90182.1; -
 DR HSSP; P15731; 10CQ.
 DR FlyBase; FBgn0035584; CG17030.
 DR InterPro; IPR000608; UQO_conjugat.
 DR Pfam; PF00179; UQ_con; 1.
 DR PRODOM; PD000461; UQO_conjugat; 1.
 DR SMART; SM00212; UBC; 1.
 DR PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
 KM LIGASE; Ubiqultin conjugation.
 SQ SEQUENCE 180 AA; 21335 MW; 78543FDB44AA0E7D CRC64;

Query Match 86.1%; Score 31; DB 5; Length 180;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LEDKON 7
 DB 23 LEDKON 28

RESULT 8
 ID Q9RH82 PRELIMINARY; PRT; 366 AA.
 AC Q9RH82;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Rod shape determining protein Mrec.
 GN MREC.
 OS Pseudomonas fluorescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=294;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MFO;
 RA Regard C., Merleau A., Guespin-Michel J.,
 RT "me locus in Pseudomonas fluorescens MFO: regulation by growth
 RT temperature";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF063934; AAF2468.1; -
 DR InterPro; IPR005223; Mrec.
 DR TIGRfams; TIGR00219; mrec; 1.
 SQ SEQUENCE 366 AA; 39203 MW; 8815E4DF05B8C56 CRC64;

Query Match 86.1%; Score 31; DB 2; Length 366;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 DB 288 ELEDKON 294

RESULT 9
 ID Q9AZC6 PRELIMINARY; PRT; 499 AA.
 AC Q9AZC6;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Putative portal protein.
 GN ORF500.
 OS Lactobacillus johnsonii prophage Lj965.
 OC Viruses.

OX NCBI_TaxID=139870;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20455575; PubMed=10998330;
 RA Desiere F., Pridmore R.D., Brusov H.;
 RT "Comparative genomics of the late gene cluster from lactobacillus
 RT phages"; 275:294-305(2000).
 RL Virology; 275:294-305(2000).
 DR EMBL; AF195900; AAK27895.1; -
 SQ SEQUENCE 499 AA; 56857 MW; 105C0832AD8E037 CRC64;

Query Match 86.1%; Score 31; DB 9; Length 499;
 Best Local Similarity 85.7%; Pred. No. 2.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 DB 471 ELEDKON 477

RESULT 10
 ID Q93Y14 PRELIMINARY; PRT; 657 AA.
 AC Q93Y14;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 73.7 kDa protein.
 OS Capsella rubella.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Capsella.
 OX NCBI_TaxID=81985;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Acarkan A., Schmidt R.H.;
 RT "Comparative analysis of the Arabidopsis thaliana and Capsella rubella
 RT genomes";
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ299420; CAC82614.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 657 AA; 73739 MW; 46E8947609147ED2 CRC64;

Query Match 86.1%; Score 31; DB 10; Length 657;
 Best Local Similarity 71.4%; Pred. No. 2.7e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 DB 409 ELEDKON 415

RESULT 11
 ID Q23544 PRELIMINARY; PRT; 674 AA.
 AC Q23544; Q23543;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Hypothetical 75.3 kDa protein.
 GN AT4G17000.
 GN Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Slekema W., Murphy G., Wambutt R., Pohl T., Terry N.,
 RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
 RA Pulidomenech P., Hatzipoulos P., Obermaier B., Duesterhoft A.,
 RA Jones J., Palmer K., Anorge W., Delseny M., Bancroft I., Mewes H.W.,
 RA Scheller C., Chalmers N.;

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; J97342; CAB10480.1; -
 DR EMBL; AL161545; CAB80971.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 674 AA; 75341 MW; F846AD8C3AC63C34 CRC64;

Query Match 86.1%; Score 31; DB 10; Length 674;
 Best Local Similarity 71.4%; Pred. No. 2.8e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 |||||
 Db 412 ELEDKON 418

RESULT 12
 O9UNV0 PRELIMINARY; PRT; 286 AA.
 AC O9UNV0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein SSO0545.
 GN SSO0545 OR ORF-C22_006.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_Taxid=2287;
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE-2133296; PubMed-11427726;
 RA She O., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aweyer M.J., Chan-Welher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Eruso G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Noc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Chidobos R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL; Y18930; CAB57754.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 286 AA; 32864 MW; 8C478AF9A8D4AB8 CRC64;

Query Match 83.3%; Score 30; DB 17; Length 286;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKQ 6
 |||||
 Db 42 ELEDKQ 47

RESULT 13
 O68333 PRELIMINARY; PRT; 293 AA.
 ID O68333;
 AC O68333;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE YOPN.
 GN YOPN.
 OS Yersinia enterocolitica.
 OC Plasmid pYve27.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 OX NCBI_Taxid=630;

RN [1]
 RP SEQUENCE OF 288-293 FROM N.A.
 RC STRAIN-W22703;
 RX MEDLINE-98190073; PubMed-9524114;
 RA Iriarte M., Sory M.P., Boland A., Boyd A.P., Mills S.D.,
 RA Lambermont I., Cornelis G.R.;
 RT "Yvea, a protein involved in control of Yop release and in
 RT translocation of Yersinia Yop effectors."
 RL EMBO J. 17:1907-1918(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-W22703;
 RA Iriarte M., Lambermont I., Kerbouch C., Cornelis G.R.;
 RT "Detailed genetic map of the pYve27 plasmid of Yersinia
 RT enterocolitica serotype O:9."
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF102990; AAD16823.1; -
 KW plasmid.
 SQ SEQUENCE 293 AA; 32925 MW; 0BACB442D103C77C CRC64;

Query Match 83.3%; Score 30; DB 2; Length 293;
 Best Local Similarity 85.7%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 |||||
 Db 96 ELEDKON 102

RESULT 14
 O93KT8 PRELIMINARY; PRT; 293 AA.
 ID O93KT8;
 AC O93KT8;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE YOPN.
 GN YOPN.
 OS Yersinia enterocolitica.
 OC Plasmid pYve8081.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 OX NCBI_Taxid=630;
 RP SEQUENCE FROM N.A.
 RC STRAIN-8081;
 RX MEDLINE-21295118; PubMed-11402007;
 RA Snellings N.J., Poppe M., Lindler L.E.;
 RT "Complete DNA Sequence of Yersinia enterocolitica Serotype O:8 Low-
 RT Calcium-Response Plasmid Reveals a New Virulence Plasmid-Associated
 RT Replicon."
 RL Infect. Immun. 69:4627-4638(2001).
 DR EMBL; AF336309; AAK69222.1; -
 KW Plasmid.
 SQ SEQUENCE 293 AA; 32642 MW; 0A3412E2CB412904 CRC64;

Query Match 83.3%; Score 30; DB 2; Length 293;
 Best Local Similarity 85.7%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 |||||
 Db 96 ELEDKON 102

RESULT 15
 O8TNU6 PRELIMINARY; PRT; 348 AA.
 ID O8TNU6;
 AC O8TNU6;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Predicted protein.

GN MA2182;
OS Methanosarcina acetivorans;
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina;
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Strange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Telamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
and physiological diversity.";
RL Genome Res. 12:532-542(2002).
DR EMBL: AE010904; AAM05579.1; .
KW Complete proteome.
SO SEQUENCE 348 AA; 40821 MW; EB64BD7EA68B4B79 CRC64;

Query Match 83.3%; Score 30; DB 17; Length 348;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKO 6
 |||||
DB 71 ELEDKO 76

Search completed: November 13, 2002, 13:16:08
Job time : 23.4894 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 8.68085 Seconds
(Without alignments)
20.336 Million cell updates/sec

Title: US-09-856-086-5
Perfect score: 32
Sequence: 1 KKVHEE 6

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCUS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	100.0	160	2 US-08-726-306A-183	Sequence 183, App
2	31	96.9	3135	1 US-08-323-1708-2	Sequence 2, Appl
3	31	96.9	3135	4 US-08-954-441-2	Sequence 2, Appl
4	29	90.6	466	4 US-09-610-401-3	Sequence 3, Appl
5	28	87.5	68	4 US-09-911-927-33	Sequence 33, Appl
6	28	87.5	68	4 US-09-911-882-33	Sequence 33, Appl
7	28	87.5	69	4 US-08-965-762-33	Sequence 33, Appl
8	28	87.5	535	2 US-08-007-107-2	Sequence 2, Appl
9	27	84.4	7	4 US-09-172-045-4	Sequence 2, Appl
10	27	84.4	448	2 US-08-878-989-2	Sequence 2, Appl
11	27	84.4	448	4 US-09-272-796-2	Sequence 2, Appl
12	27	84.4	494	4 US-09-126-420A-26	Sequence 26, Appl
13	27	84.4	598	4 US-09-344-700-4	Sequence 4, Appl
14	27	84.4	549	4 US-08-325-071-61	Sequence 61, Appl
15	27	84.4	549	4 US-08-461-004A-61	Sequence 61, Appl
16	27	84.4	620	1 US-08-325-071-65	Sequence 65, Appl
17	27	84.4	620	4 US-08-461-004A-65	Sequence 65, Appl
18	27	84.4	620	4 US-08-325-071-56	Sequence 56, Appl
19	27	84.4	650	1 US-08-325-071-59	Sequence 59, Appl
20	27	84.4	650	1 US-08-325-071-63	Sequence 63, Appl
21	27	84.4	650	1 US-08-325-071-67	Sequence 67, Appl
22	27	84.4	650	4 US-08-461-004A-56	Sequence 56, Appl
23	27	84.4	650	4 US-08-461-004A-59	Sequence 59, Appl
24	27	84.4	650	4 US-08-461-004A-63	Sequence 63, Appl
25	27	84.4	650	4 US-08-461-004A-67	Sequence 67, Appl
26	27	84.4	688	1 US-08-325-071-57	Sequence 57, Appl
27	27	84.4	688	4 US-08-461-004A-57	Sequence 57, Appl

28	26	81.2	13	1 US-07-987-272A-6	Sequence 6, Appl
29	26	81.2	46	1 US-08-200-016-3	Sequence 3, Appl
30	26	81.2	93	1 US-07-987-272A-7	Sequence 7, Appl
31	26	81.2	93	1 US-07-987-272A-16	Sequence 16, Appl
32	26	81.2	93	1 US-08-385-241-1	Sequence 21, Appl
33	26	81.2	283	4 US-09-145-828A-22	Sequence 22, Appl
34	26	81.2	284	2 US-09-014-969-21	Sequence 21, Appl
35	26	81.2	303	6 5340934-13	Patent No. 5340934
36	26	81.2	364	3 US-08-776-900C-23	Sequence 23, Appl
37	26	81.2	364	4 US-09-268-195C-23	Sequence 23, Appl
38	26	81.2	376	4 US-08-983-502-11	Sequence 11, Appl
39	26	81.2	376	5 PCT-US96-10521-11	Sequence 11, Appl
40	26	81.2	418	1 US-08-224-930-4	Sequence 4, Appl
41	26	81.2	418	4 US-08-908-436-6	Sequence 6, Appl
42	26	81.2	418	4 US-09-561-756-18	Sequence 18, Appl
43	26	81.2	418	4 US-09-227-721-18	Sequence 18, Appl
44	26	81.2	420	2 US-08-592-126-142	Sequence 142, App
45	26	81.2	466	4 US-09-610-401-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-726-306A-183
Sequence 183, Application US/08726306A
Patent No. 5958584
GENERAL INFORMATION:
APPLICANT: van Leeuwen, Frederik Willem
APPLICANT: Burbach, Johannes Peter Henri
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:
ADDRESS: Banner & Witcoff, Ltd.
STREET: 1 Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,306A
FILING DATE: 02-Oct-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080.4
FILING DATE: 02-Oct-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/009,832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 183:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-726-306A-183
Query Match 100.0%; Score 32; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
|||||
DB 10 KKVHEE 15

RESULT 2
US-08-323-1708-2
Sequence 2, Application US/083231708
Patent No. 5733772
GENERAL INFORMATION:
APPLICANT: Williamson, Kim C.
APPLICANT: Kaslow, David C.
TITLE OF INVENTION: Cloning and Expression of Plasmodium
TITLE OF INVENTION: falciparum Transmission-Blocking Target Antigen, Pf230
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,1708
FILING DATE: 13-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Quine, Jonathan A.
REGISTRATION NUMBER: P-41,261
REFERENCE/DOCKET NUMBER: 015280-113100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-323-1708-2
Query Match 96.9%; Score 31; DB 1; Length 3135;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKVHEE 6
|||||
DB 2556 KKVHEE 2561

RESULT 3
US-08-954-441-2
Sequence 2, Application US/08954441
Patent No. 6316000
GENERAL INFORMATION:
APPLICANT: Williamson, Kim C.
APPLICANT: Kaslow, David C.
TITLE OF INVENTION: Cloning and Expression of Plasmodium
TITLE OF INVENTION: falciparum Transmission-Blocking Target Antigen, Pf230
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco

STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,441
FILING DATE: 20-OCT-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,170
FILING DATE: 13-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Einhorn, Gregory P.
REGISTRATION NUMBER: 38,440
REFERENCE/DOCKET NUMBER: 015280-113110US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-954-441-2
Query Match 96.9%; Score 31; DB 4; Length 3135;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKVHEE 6
|||||
DB 2556 KKVHEE 2561

RESULT 4
US-09-610-401-3
Sequence 3, Application US/09610401
Patent No. 6417336
GENERAL INFORMATION:
APPLICANT: MORISHIMA, No. 6417336unltd,
APPLICANT: NAKANISHI, Keiko,
APPLICANT: SHIBATA, Takehiko
TITLE OF INVENTION: Antibody against cleavage product of vimentin
FILE REFERENCE: 522,1004
CURRENT APPLICATION NUMBER: US/09/610,401
CURRENT FILING DATE: 2000-07-05
PRIOR APPLICATION NUMBER: JP 11-193235
PRIOR FILING DATE: 1999-07-07
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-09-610-401-3
Query Match 90.6%; Score 29; DB 4; Length 466;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKVHEE 6
|||||
DB 235 KKVHEE 240

```
RESULT 5
US-09-911-927-33
; Sequence 33, Application US/09911927
; Patent No. 6461826
; ORGANISM: Aspergillus nidulans
; GENERAL INFORMATION:
; APPLICANT: Koltin, Yigal
; APPLICANT: Gavrilas, Victoria
; TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
; FILE REFERENCE: 06286-062003
; CURRENT APPLICATION NUMBER: US/09/911,927
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 08/965,762
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-09-911-927-33

Query Match
Best Local Similarity 87.5%; Score 28; DB 4; Length 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
|||||
DB 8 KKVHEK 13

RESULT 6
US-09-911-882-33
; Sequence 33, Application US/09911882
; Patent No. 6465198
; ORGANISM: Aspergillus nidulans
; GENERAL INFORMATION:
; APPLICANT: Koltin, Yigal
; APPLICANT: Gavrilas, Victoria
; TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
; FILE REFERENCE: 06286-062004
; CURRENT APPLICATION NUMBER: US/09/911,882
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 08/965,762
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-09-911-882-33

Query Match
Best Local Similarity 87.5%; Score 28; DB 4; Length 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
|||||
DB 8 KKVHEK 13

RESULT 7
US-08-965-762-33
; Sequence 33, Application US/08965762
; Patent No. 6280863
; ORGANISM: Aspergillus nidulans
; GENERAL INFORMATION:
; APPLICANT: Koltin, Yigal
; APPLICANT: Gavrilas, Victoria
; TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
; FILE REFERENCE: 07334/062001
; CURRENT APPLICATION NUMBER: US/08/965,762
; CURRENT FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 33
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-08-965-762-33

Query Match
Best Local Similarity 87.5%; Score 28; DB 4; Length 69;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
|||||
DB 9 KKVHEK 14

RESULT 8
US-08-007-107-2
; Sequence 2, Application US/08007107
; Patent No. 5837545
; ORGANISM: Aspergillus nidulans
; GENERAL INFORMATION:
; APPLICANT: Guy, Charles L.
; APPLICANT: Haskell, Dale W.
; APPLICANT: Hotig, Andrea
; APPLICANT: Neven, Lisa
; TITLE OF INVENTION: No. 5837545el Genes, Polypeptides, and
; TITLE OF INVENTION: Compositions for Cold Tolerance and Drought Resistance in
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/007,107
; FILING DATE: 19930121
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF/S&S-109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 535 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-007-107-2

Query Match
Best Local Similarity 87.5%; Score 28; DB 2; Length 535;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKVHEE 6
|||||
DB 525 KKVHEE 530

RESULT 9
US-09-172-045-4
; Sequence 4, Application US/09172045
; Patent No. 6277594
```

GENERAL INFORMATION:
APPLICANT: Mikoshiba, Katsuhiko
APPLICANT: Aruga, Jun
APPLICANT: Nagai, Takeharu
APPLICANT: Nakata, Katsunori
TITLE OF INVENTION: Neurogenesis Inducing Gene
FILE REFERENCE: HIRAKI-03497
CURRENT FILING DATE: 1998-10-08
EARLIER APPLICATION NUMBER: JP98/86979
EARLIER FILING DATE: 1998-03-31
EARLIER APPLICATION NUMBER: JP98/121456
EARLIER FILING DATE: 1998-04-30
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Designed
OTHER INFORMATION: peptide based on amino acid sequence of zinc
OTHER INFORMATION: finger motif of mouse zic gene family.
US-09-172-043-4

Query Match 84.4%; Score 27; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 KYHEE 6
|||||
DB 3 KYHEE 7

RESULT 10
US-08-878-989-2
Sequence 2, Application US/08678989
Patent No. 5885803
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lai, Preeti
APPLICANT: Goll, Surya K.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878, 989
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TBLYN0T01
CLONE: 40194
US-08-878-989-2

Query Match 84.4%; Score 27; DB 2; Length 448;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KYHEE 6
|||||
DB 302 KYHEE 307

RESULT 11
US-09-272-796-2
Sequence 2, Application US/09272796
Patent No. 6207148
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lai, Preeti
APPLICANT: Goll, Surya K.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/272,796
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/878, 989
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: TBLYN0T01

CLONE: 40194
US-09-272-796-2

Query Match 84.4%; Score 27; DB 4; Length 448;
Best Local Similarity 83.3%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKVHEE 6
|||||
DB 302 KKVHSE 307

RESULT 12
US-09-126-420A-26

Sequence 26, Application US/09126420A
Patent No. 6376753

GENERAL INFORMATION:

APPLICANT: BATAARD, YANNICK

APPLICANT: ROBINNEAU, TIBURCE

APPLICANT: DURST, FRANCIS

APPLICANT: MERCK-REICHAERT, DANIELE

APPLICANT: DIDIERJEAN, LUC

TITLE OF INVENTION: PURIFIED CYTOCHROME P450 CYP76B1 FROM HELIANTHUS

TITLE OF INVENTION: TUBEROSUS AND ITS APPLICATIONS AS BIOCATALYST IN

TITLE OF INVENTION: PARTICULAR FOR THE DEGRADATION OF ENVIRONMENTAL

FILE REFERENCE: 03715.0032

CURRENT FILING DATE: 1998-07-30

PRIOR APPLICATION NUMBER: US/09/126.420A

PRIOR FILING DATE: 1997-07-31

NUMBER OF SEQ ID NOS: 27

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 26

LENGTH: 494

TYPE: PRT

ORGANISM: Rattus norvegicus

US-09-126-420A-26

Query Match 84.4%; Score 27; DB 4; Length 494;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KKVHEE 6
|||||
DB 326 KKVHSE 330

RESULT 13
US-09-344-700-4

Sequence 4, Application US/09344700
Patent No. 6265194

GENERAL INFORMATION:

APPLICANT: Nezu, Jun-ichi

TITLE OF INVENTION: NOVEL SERINE-THREONINE KINASE GENE

FILE REFERENCE: 06501/033001

CURRENT FILING DATE: 1999-06-25

EARLIER FILING DATE: 1997-12-25

EARLIER FILING DATE: 1996-12-27

NUMBER OF SEQ ID NOS: 37

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4

LENGTH: 508

TYPE: PRT

ORGANISM: Homo sapiens

US-09-344-700-4

Query Match 84.4%; Score 27; DB 4; Length 508;
Best Local Similarity 83.3%; Pred. No. 4.9e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 KKVHEE 6
|||||
DB 362 KKVHSE 367

RESULT 14
US-08-325-071-61

Sequence 61, Application US/08325071
Patent No. 5587311

GENERAL INFORMATION:

APPLICANT: COBON, Stewart Gary

APPLICANT: MOORE, Joanna Terry

APPLICANT: JOHNSON, Law Anthony York

APPLICANT: WILLADSEN, Peter

APPLICANT: KEMP, David Harold

APPLICANT: SRISKANTHA, Alagacone

APPLICANT: RIDING, George Alfred

APPLICANT: RAND, Keith No. 3587311man

TITLE OF INVENTION: DNA Encoding A Cell Membrane

TITLE OF INVENTION: Glycoprotein Of A Tick Gut

NUMBER OF SEQUENCES: 71

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 3000 K Street, N.W.

CITY: Washington, D.C.

COUNTRY: USA

ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/325.071

FILING DATE: 14-OCT-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/062,109

FILING DATE: 17-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/926,368

FILING DATE: 07-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/242,196

FILING DATE: 06-JUL-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/AU87/00401

FILING DATE: 27-NOV-1987

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU P14912

FILING DATE: 16-OCT-1987

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU P12570

FILING DATE: 19-JUN-1987

PRIOR APPLICATION DATA:

APPLICATION NUMBER: AU PH9196

FILING DATE: 27-NOV-1986

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 60042/111 B1AU

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202 672 5300

TELEFAX: 202 672 5399

TELEX: 904136

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:

LENGTH: 549 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-325-071-61

Query Match 84.4% Score 27; DB 1; Length 549;
Best Local Similarity 83.3% Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KKVHEE 6

Db 217 KKVHEE 222

RESULT 15
US-08-461-004A-61
Sequence 61, Application US/08461004A
Patent No. 6235283
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISANTHA, Alagacone
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 6235283man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
TITLE OF INVENTION: Glycoprotein Of A Tick Gut
NUMBER OF SEQUENCES: 71
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,004A
FILING DATE: 04-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/325,071
FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI4912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI2570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 60042/152
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 904136
INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:
LENGTH: 549 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-004A-61

Query Match 84.4% Score 27; DB 4; Length 549;
Best Local Similarity 83.3% Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KKVHEE 6
Db 217 KKVHEE 222

Search completed: November 13, 2002, 13:18:07
Job time : 8.68085 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:18:18 ; Search time 4.08511 Seconds
(without alignments)
22.121 Million cell updates/sec

Title: US-09-856-086-5

Perfect score: 32

Sequence: 1 KKVHEE 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCIT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCIT_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	90.6	466	10 US-09-923-779-155	Sequence 155, App
2	29	90.6	466	12 US-10-152-647-3	Sequence 3, Appl
3	28	87.5	68	10 US-09-911-888-33	Sequence 33, Appl
4	28	87.5	134	10 US-09-864-761-34684	Sequence 34684, A
5	28	87.5	398	10 US-09-729-674-146	Sequence 146, App
6	28	87.5	400	10 US-09-881-752A-206	Sequence 206, App
7	27	84.4	39	10 US-09-864-761-42664	Sequence 42664, A
8	27	84.4	62	10 US-09-864-761-47501	Sequence 47501, A
9	26	81.2	64	10 US-09-864-761-39889	Sequence 39889, A
10	26	81.2	76	10 US-09-864-761-34168	Sequence 34168, A
11	26	81.2	82	10 US-09-864-761-46614	Sequence 46614, A
12	26	81.2	104	10 US-09-864-761-37686	Sequence 37686, A
13	26	81.2	157	10 US-09-741-669-427	Sequence 427, App
14	26	81.2	283	10 US-09-903-456-29	Sequence 29, Appl
15	26	81.2	336	10 US-09-764-898-281	Sequence 281, App
16	26	81.2	418	10 US-09-934-867-18	Sequence 18, Appl
17	26	81.2	435	10 US-09-866-582-33	Sequence 33, Appl
18	26	81.2	466	12 US-10-152-647-4	Sequence 4, Appl
19	26	81.2	600	10 US-09-975-901-2	Sequence 2, Appl

20	26	81.2	694	9 US-09-944-413-55	Sequence 55, Appl
21	26	81.2	694	9 US-09-944-403-55	Sequence 55, Appl
22	26	81.2	694	9 US-09-866-028-55	Sequence 55, Appl
23	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
24	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
25	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
26	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
27	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
28	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
29	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
30	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
31	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
32	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
33	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
34	26	81.2	694	10 US-09-944-449-55	Sequence 55, Appl
35	26	81.2	725	10 US-09-815-242-5125	Sequence 5125, Ap
36	26	81.2	758	10 US-09-903-023-2	Sequence 2, Appl
37	26	81.2	758	10 US-09-859-604-2	Sequence 2, Appl
38	26	81.2	758	10 US-09-903-063-2	Sequence 2, Appl
39	26	81.2	758	10 US-09-903-216-2	Sequence 2, Appl
40	26	81.2	758	10 US-09-903-199-2	Sequence 2, Appl
41	26	81.2	758	10 US-09-903-023-2	Sequence 2, Appl
42	26	81.2	797	10 US-09-815-242-4882	Sequence 4882, Ap
43	26	81.2	798	10 US-09-815-242-10764	Sequence 10764, A
44	26	81.2	942	10 US-09-828-423-1	Sequence 1, Appl
45	26	81.2	1352	10 US-09-736-969A-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-09-923-779-155
Sequence 155, Application US/09923779
Patent No. US20020076721A1
GENERAL INFORMATION:
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Kalos, Michael D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.553
CURRENT APPLICATION NUMBER: US/09/923, 779
CURRENT FILING DATE: 2001-08-06
NUMBER OF SEQ ID NOS: 155
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 155
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-09-923-779-155

Query Match 90.6%; Score 29; DB 10; Length 466;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
DB 235 KKVHEE 240

RESULT 2
US-10-152-647-3
Sequence 3, Application US/10152647
Patent No. US20020137110A1
GENERAL INFORMATION:
APPLICANT: MORISHIMA, NO. US20020137110A1uh1ro,
APPLICANT: NAKANISHI, Keko,
APPLICANT: SHIBATA, Takehiko
TITLE OF INVENTION: Antibody against cleavage product of vimentin
FILE REFERENCE: 522.1004
CURRENT APPLICATION NUMBER: US/10/152,647
CURRENT FILING DATE: 2002-05-21

PRIOR APPLICATION NUMBER: US/09/610,401
PRIOR FILING DATE: 2000-07-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-10-152-647-3

Query Match 90.6%; Score 29; DB 12; Length 466;
Best Local Similarity 83.3%; Pred. No. 66;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
DB 235 KKLHEE 240

RESULT 3
US-09-911-888-33
Sequence 33, Application US/09911888
Patent No. US20020119509A1
GENERAL INFORMATION:
APPLICANT: Koltin, Yigal
TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
FILE REFERENCE: 06286-062002
CURRENT APPLICATION NUMBER: US/09/911,888
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 08/965,762
PRIOR FILING DATE: 1997-11-07
NUMBER OF SEQ ID NOS: 35
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 68
TYPE: PRT
ORGANISM: Aspergillus nidulans
US-09-911-888-33

Query Match 87.5%; Score 28; DB 10; Length 68;
Best Local Similarity 83.3%; Pred. No. 14;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
DB 8 KKVHEK 13

RESULT 4
US-09-864-761-34684
Sequence 34684, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aegm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34684
LENGTH: 134
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC010680.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
OTHER INFORMATION: EST_HUMAN HIT: AA196127.1, EVALUE 2.00e-25
OTHER INFORMATION: SWISSPROT HIT: Q28820, EVALUE 5.60e-01
US-09-864-761-34684

Query Match 87.5%; Score 28; DB 10; Length 134;
Best Local Similarity 83.3%; Pred. No. 28;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
DB 86 KKVHEE 91

RESULT 5
US-09-729-674-146
Sequence 146, Application US/09729674
Patent No. US20010039335A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Werberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Agostino, Michael J.
APPLICANT: Steininger II, Robert J.
APPLICANT: Spaulding, Vikki
APPLICANT: Wood, Gordon G.
APPLICANT: Clark, Hilary
APPLICANT: Fechtel, Kim
APPLICANT: Genetics Institute, Inc.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: 6055-64X
CURRENT APPLICATION NUMBER: US/09/729,674

;; CURRENT FILING DATE: 2000-12-04
;; PRIOR APPLICATION NUMBER: 09/539,330
;; PRIOR FILING DATE: 2000-03-30
;; NUMBER OF SEQ ID NOS: 283
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 146
;; LENGTH: 398
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: UNSURE
;; LOCATION: (379)
US-09-729-674-146

Query Match 87.5%; Score 28; DB 10; Length 398;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
:|||||
DB 230 KKVHEE 235

RESULT 6
US-09-881-752A-206
;; Sequence 206, Application US/09881752A
;; Patent No. US20020115078A1
;; GENERAL INFORMATION:
;; APPLICANT: Kleantous, Harold
;; APPLICANT: Al-Garawi, Amal
;; APPLICANT: Miller, Charles
;; APPLICANT: Tomb, Jean-Francois
;; APPLICANT: Oomen, Raymond P.

;; TITLE OF INVENTION: Identification of Polynucleotides
;; TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
;; FILE REFERENCE: 06132/041002
;; CURRENT APPLICATION NUMBER: US/09/881,752A
;; CURRENT FILING DATE: 2001-06-15
;; PRIOR APPLICATION NUMBER: US 08/833,457
;; PRIOR FILING DATE: 1997-04-01
;; NUMBER OF SEQ ID NOS: 370
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 206
;; LENGTH: 400
;; TYPE: PRT
;; ORGANISM: Helicobacter pylori
US-09-881-752A-206

Query Match 87.5%; Score 28; DB 10; Length 400;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 KKVHEE 6
:|||||
DB 35 KKVHEE 40

RESULT 7
US-09-864-761-42664
;; Sequence 42664, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aecmiga-x-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312

;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 42664
;; LENGTH: 39
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AP00246.1
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.61
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.67
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.7
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.86
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.65
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.7
;; OTHER INFORMATION: EST_HUMAN HIT: AA233606.1, EVALUE 3.00e-15
;; OTHER INFORMATION: SWISSPROT HIT: Q06010, EVALUE 3.00e-16
US-09-864-761-42664

Query Match 84.4%; Score 27; DB 10; Length 39;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
:|||||
DB 16 KKVHEE 21

RESULT 8
US-09-864-761-47501
;; Sequence 47501, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng

```
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT FILING DATE: 2001-05-23
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 47501
;; LENGTH: 62
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC008962.4
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 9.9
;; OTHER INFORMATION: SWISSPROT HIT: Q16966, EVALU6 6.00e-32
;; OTHER INFORMATION: EST_HUMAN HIT: AW577399.1, EVALU6 5.00e-25
US-09-864-761-47501
Query Match      84.4%; Score 27; DB 10; Length 62;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      2 KVEEE 6
Db      1 KVEEE 5
RESULT 9
US-09-864-761-39889
;; Sequence 39889, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
```

```
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT FILING DATE: 2001-05-23
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 39889
;; LENGTH: 64
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC004775.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.9
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.7
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.5
;; OTHER INFORMATION: SWISSPROT HIT: Q92599, EVALU6 2.00e-19
;; OTHER INFORMATION: EST_HUMAN HIT: BE748158.1, EVALU6 2.00e-18
US-09-864-761-39889
Query Match      81.2%; Score 26; DB 10; Length 64;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY      1 KVEEE 6
Db      9 KRVHE 14
RESULT 10
US-09-864-761-34168
;; Sequence 34168, Application US/09864761
;; Patent No. US20020048763A1
```

```
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34168
LENGTH: 76
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000049.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.9
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
OTHER INFORMATION: SWISSPROT HIT: Q15811, EVALUATE 9.00e-25
OTHER INFORMATION: EST_HUMAN HIT: AA313515.1, EVALUATE 5.00e-24
US-09-864-761-34168

Query Match      81.2%; Score 26; DB 10; Length 76;
Best Local Similarity 66.7%; Pred. No. 40;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
DB      23 RKLHEE 28      :|||
RESULT 11
US-09-864-761-46614
Sequence 46614, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecmca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 46614
LENGTH: 82
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000311.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.84
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.61
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.63
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.82
OTHER INFORMATION: EST_HUMAN HIT: AW950164.1, EVALUATE 9.00e-24
OTHER INFORMATION: SWISSPROT HIT: Q15811, EVALUATE 1.00e-24
US-09-864-761-46614

Query Match      81.2%; Score 26; DB 10; Length 82;
```

Best Local Similarity 66.7%; Pred. No. 43;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHE 6
11:111
Db 23 KKLHE 28

RESULT 12
US-09-864-761-37686
; Sequence 37686, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1.
; SEQ ID NO 37686
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC012443.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL - 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 1.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL - 3.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 2.9
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 2.2

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 0.98
; OTHER INFORMATION: EST_HUMAN HIT: AL042120.1, EVALUATE 1.00e-33
; OTHER INFORMATION: SWISSPROT HIT: O54786, EVALUATE 1.30e+00
US-09-864-761-37686

Query Match 81.2%; Score 26; DB 10; Length 104;
Best Local Similarity 80.0%; Pred. No. 55;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHE 5
11:11
Db 32 KKLHE 36

RESULT 13
US-09-741-669-427
; Sequence 427, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; FILE REFERENCE: ELITRA.009A
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 427
; LENGTH: 157
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-427

Query Match 81.2%; Score 26; DB 10; Length 157;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 KKVHE 6
11:111
Db 141 KKLHE 146

RESULT 14
US-09-903-456-29
; Sequence 29, Application US/09903456
; Patent No. US20020138874A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Mukerji, Pradip
; APPLICANT: Leonard, Amanda Eun-yeong
; APPLICANT: Huang, Yung-Sheng
; APPLICANT: Pereira, Suzette L.
; TITLE OF INVENTION: ELONGASE GENES AND USES THEREOF
; FILE REFERENCE: 6407 US P3
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 09/624,670
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: US 09/379,095
; PRIOR FILING DATE: 1999-08-23
; PRIOR APPLICATION NUMBER: US 09/145,828
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 283
; TYPE: PRT


```

; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Potential Mammalian Elongase
; NAME/KEY: VARIANT
; LOCATION: (282)...(282)
; OTHER INFORMATION: Xaa - Unknown or Other at position 282
US-09-903-456-29

```

```

Query Match      81.2%; Score 26; DB 10; Length 283;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2 KVVHE 6
        1:111
Db      33 KVVHE 37

```

```

RESULT 15
US-09-764-898-281
; Sequence 281, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ201
; CURRENT APPLICATION NUMBER: US/09/764,898
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 281
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (168)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (186)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (299)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (318)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (329)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-898-281

```

```

Query Match      81.2%; Score 26; DB 10; Length 336;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      1 KVVHE 6
        1:111
Db      270 KVVHE 275

```

Search completed: November 13, 2002, 13:40:19
 Job time : 5.22796 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:16:18 : Search time 120.383 Seconds
(without alignments)
32.134 Million cell updates/sec

Title: US-09-856-086-5

Perfect score: 32

Sequence: 1 KKVHEE 6

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	32	100.0	448	21	US-09-791-537-94919
2	32	100.0	450	25	US-10-108-260A-4090
3	32	100.0	455	21	US-09-791-537-32570
4	32	100.0	461	21	US-09-791-537-94920
5	32	100.0	462	21	US-09-791-537-94921
6	32	100.0	463	21	US-09-791-537-73564

Query	Match	Score	DB	Length	Score	DB	Length
1 KKVHEE 6	100.0%	32	21	448	100.0%	32	448
223 KKVHEE 228	100.0%	32	21	448	100.0%	32	448

ALIGNMENTS

RESULT 1
US-09-791-537-94919
Sequence 94919, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 94919
LENGTH: 448
TYPE: PRT
ORGANISM: Gallus gallus
US-09-791-537-94919

Query Match: 100.0%; Score 32; DB 21; Length 448;
Best Local Similarity: 100.0%; Pred. No. 9.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
US-10-108-260A-4090
; Sequence 4090, Application US/10108260A
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4090
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4090

Query Match
Best Local Similarity 100.0%; Score 32; DB 25; Length 450;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
|||||

DB 220 KKVHEE 225

RESULT 3
US-09-791-537-32570
; Sequence 32570, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32570
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Scyllorhinus stellaris
US-09-791-537-32570

Query Match
Best Local Similarity 100.0%; Score 32; DB 21; Length 455;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
|||||

DB 224 KKVHEE 229

RESULT 4
US-09-791-537-94920
; Sequence 94920, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 94920
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Bos taurus

US-09-791-537-94920

Query Match
Best Local Similarity 100.0%; Score 32; DB 21; Length 461;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
|||||

DB 231 KKVHEE 236

RESULT 5
US-09-791-537-94921
; Sequence 94921, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY M
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 94921
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-791-537-94921

Query Match
Best Local Similarity 100.0%; Score 32; DB 21; Length 462;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
|||||

DB 232 KKVHEE 237

RESULT 6
US-09-791-537-73564
; Sequence 73564, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blonomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY M
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73564
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-791-537-73564

Query Match
Best Local Similarity 100.0%; Score 32; DB 21; Length 463;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
|||||

DB 231 KKVHEE 236

RESULT 7
US-09-538-092-948
; Sequence 948, Application US/09538092
; GENERAL INFORMATION:

```
;; APPLICANT: Glot, Loic
;; APPLICANT: Mansfield, Traci A.
;; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;; FILE REFERENCE: 15966-542
;; CURRENT APPLICATION NUMBER: US/09/538,092
;; CURRENT FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: 60/127,352
;; PRIOR FILING DATE: 1999-04-01
;; PRIOR APPLICATION NUMBER: 60/178,965
;; PRIOR FILING DATE: 2000-02-01
;; NUMBER OF SEQ ID NOS: 1387
;; SOFTWARE: Curataseqformatter Version 0.9
;; SEQ ID NO 948
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc.feature
;; LOCATION: (0)..(0)
;; OTHER INFORMATION: Polypeptide Accession Number P17661
US-09-538-092-948

Query Match          100.0%; Score 32; DB 19; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
    |||||
Db 239 KKVHEE 244

RESULT 8
US-09-791-537-62363
;; Sequence 62363, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomolx, Inc.
;; APPLICANT: Debe, Derek
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 62363
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Rattus norvegicus
US-09-791-537-62363

Query Match          100.0%; Score 32; DB 21; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
    |||||
Db 239 KKVHEE 244

RESULT 9
US-09-791-537-73566
;; Sequence 73566, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomolx, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
```

```
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 73566
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Mesocricetus auratus
US-09-791-537-73566

Query Match          100.0%; Score 32; DB 21; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
    |||||
Db 239 KKVHEE 244

RESULT 10
US-09-791-537-132613
;; Sequence 132613, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomolx, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 132613
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-09-791-537-132613

Query Match          100.0%; Score 32; DB 21; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
    |||||
Db 239 KKVHEE 244

RESULT 11
US-60-389-987-213
;; Sequence 213, Application US/60389987
;; GENERAL INFORMATION:
;; APPLICANT: Ghosh, Soumitra S.
;; APPLICANT: Fahy, Colin D.
;; APPLICANT: Zhang, Bing
;; APPLICANT: Gibson, Bradford W.
;; APPLICANT: Taylor, Steven W.
;; APPLICANT: Glenn, Gary M.
;; APPLICANT: Warnock, Dale E.
;; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
;; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
;; FILE REFERENCE: 660086.465P2
;; CURRENT APPLICATION NUMBER: US/60/389,987
;; CURRENT FILING DATE: 2002-06-17
;; NUMBER OF SEQ ID NOS: 3025
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 213
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-389-987-213

Query Match          100.0%; Score 32; DB 27; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 KKVHEE 6
|||||
Db 239 KKVHEE 244

RESULT 12
US-60-412-418-213
Sequence 213, Application US/60412418
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Eoin D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Marnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
FILE REFERENCE: 660088.465P3
CURRENT APPLICATION NUMBER: US/60/412.418
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ ID NOS: 3025
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 213
LENGTH: 469
TYPE: PRT
ORGANISM: Homo sapiens
US-60-412-418-213

Query Match 100.0%; Score 32; DB 27; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
|||||
Db 239 KKVHEE 244

RESULT 13
US-09-791-537-46962
Sequence 46962, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Dede, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 46962
LENGTH: 470
TYPE: PRT
ORGANISM: Bos taurus
US-09-791-537-46962

Query Match 100.0%; Score 32; DB 21; Length 470;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
|||||
Db 240 KKVHEE 245

RESULT 14
US-09-791-537-62106
Sequence 62106, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Dede, Derek

APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 62106
LENGTH: 470
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-62106

Query Match 100.0%; Score 32; DB 21; Length 470;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
|||||
Db 240 KKVHEE 245

RESULT 15
US-09-791-537-65178
Sequence 65178, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Dede, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791.537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 65178
LENGTH: 470
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-65178

Query Match 100.0%; Score 32; DB 21; Length 470;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
|||||
Db 240 KKVHEE 245

Search completed: November 13, 2002, 13:39:18
Job time : 121.383 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: November 13, 2002, 13:17:03 ; Search time 2.42553 Seconds
(without alignments)
29.431 Million cell updates/sec

Title: US-09-856-086-5
Perfect score: 32
Sequence: 1 KKVHEE 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41632 seqs, 11897606 residues

Total number of hits satisfying chosen parameters: 41632

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCR_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	6	5	US-09-856-086-5
2	32	100.0	7	5	US-09-856-086-2
3	29	90.6	76	5	US-09-513-999C-7614
4	27	84.4	135	6	US-10-141-531-23
5	27	84.4	315	6	US-10-141-531-66
6	27	84.4	316	6	US-10-141-531-85
7	27	84.4	316	6	US-10-141-531-94
8	27	84.4	316	6	US-10-141-531-103
9	27	84.4	316	6	US-10-141-531-112
10	27	84.4	316	6	US-10-141-531-121
11	27	84.4	316	6	US-10-141-531-130
12	27	84.4	316	6	US-10-141-531-139
13	27	84.4	316	6	US-10-141-531-148
14	27	84.4	316	6	US-10-141-531-157
15	27	84.4	316	6	US-10-141-531-166
16	27	84.4	316	6	US-10-141-531-175
17	27	84.4	316	6	US-10-141-531-184
18	27	84.4	316	6	US-10-141-531-193
19	27	84.4	316	6	US-10-141-531-202
20	27	84.4	561	6	US-10-136-728-70
21	27	84.4	561	6	US-10-136-728-72
22	27	84.4	600	6	US-10-264-237-1886
23	26	81.2	151	6	US-10-264-237-2018
24	26	81.2	234	6	US-10-154-678-20
25	26	81.2	372	6	US-10-094-507-35
26	26	81.2	443	6	US-10-113-709A-2

27	26	81.2	694	6	US-10-125-923A-60	Sequence 60, Appl
28	26	81.2	804	6	US-10-092-411A-5218	Sequence 5218, Ap
29	26	81.2	1713	6	US-10-270-333-177	Sequence 177, Ap
30	25	78.1	184	6	US-10-264-237-2744	Sequence 2744, Ap
31	25	78.1	187	6	US-10-281-673-6	Sequence 6, Appl
32	25	78.1	693	6	US-10-264-237-2727	Sequence 2727, Ap
33	24	75.0	55	6	US-10-194-746-7	Sequence 7, Appl
34	24	75.0	66	5	US-09-513-999C-5468	Sequence 5468, Ap
35	24	75.0	71	5	US-09-513-999C-6941	Sequence 6941, Ap
36	24	75.0	121	5	US-09-513-999C-6774	Sequence 6774, Ap
37	24	75.0	149	6	US-10-092-411A-4154	Sequence 4154, Ap
38	24	75.0	154	6	US-10-092-411A-4529	Sequence 4529, Ap
39	24	75.0	164	6	US-10-092-411A-4630	Sequence 4630, Ap
40	24	75.0	166	6	US-10-267-718-7	Sequence 7, Appl
41	24	75.0	171	6	US-10-092-411A-5286	Sequence 5286, Ap
42	24	75.0	225	6	US-10-092-411A-5388	Sequence 5388, Ap
43	24	75.0	280	6	US-10-125-923A-30	Sequence 30, Appl
44	24	75.0	409	6	US-10-092-411A-3790	Sequence 3790, Ap
45	24	75.0	461	6	US-10-092-411A-3821	Sequence 3821, Ap

ALIGNMENTS

RESULT 1
US-09-856-086-5
; Sequence 5, Application US/09856086
; GENERAL INFORMATION:
; APPLICANT: EBRINGER, ALAN
; TITLE OF INVENTION: DIAGNOSIS OF DEMYELINATING OR SPONGIFORM DISEASE
; FILE REFERENCE: 78104.040
; CURRENT APPLICATION NUMBER: US/09/856, 086
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens, Bos
US-09-856-086-5.

Query Match 100.0%; Score 32; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.8e+04;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
DB 1 KKVHEE 6

RESULT 2
US-09-856-086-2
; Sequence 2, Application US/09856086
; GENERAL INFORMATION:
; APPLICANT: EBRINGER, ALAN
; TITLE OF INVENTION: DIAGNOSIS OF DEMYELINATING OR SPONGIFORM DISEASE
; FILE REFERENCE: 78104.040
; CURRENT APPLICATION NUMBER: US/09/856, 086
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens, Bos
US-09-856-086-2

Query Match 100.0%; Score 32; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.8e+04;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKVHEE 6
|||||||

Db 2 KKVHEE 7

RESULT 3
US-09-513-999C-7614

Sequence 7614, Application US/09513999C
GENERAL INFORMATION:

APPLICANT: Dumas Mline Edwards, J.B.

APPLICANT: Ductert, A.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

FILE REFERENCE: 59.US2.REG

CURRENT APPLICATION NUMBER: US/09/513, 999C

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/122,487

PRIOR FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 36681

SOFTWARE: Patent.Pm

SEQ ID NO 7614

LENGTH: 76

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: UNSURE

LOCATION: 3

OTHER INFORMATION: Xaa-Ala or Gly

FEATURE:

NAME/KEY: UNSURE

LOCATION: 4

OTHER INFORMATION: Xaa-Ile or Met

FEATURE:

NAME/KEY: UNSURE

LOCATION: 5

OTHER INFORMATION: Xaa-Asp or Asn

FEATURE:

NAME/KEY: UNSURE

LOCATION: 9

OTHER INFORMATION: Xaa-Phe or Val

FEATURE:

NAME/KEY: UNSURE

LOCATION: 46

OTHER INFORMATION: Xaa-Cys or Phe or His or Ile or Leu or Asn or Pro or Arg or Ser

US-09-513-999C-7614

Query Match 90.6%; Score 29; DB 5; Length 76;

Best Local Similarity 83.3%; Pred. No. 6.8;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6

Db 62 KKLHEE 67

RESULT 4
US-10-141-531-23

Sequence 23, Application US/10141531

GENERAL INFORMATION:

APPLICANT: Briggs, Steven P.

APPLICANT: Dalmia, Bipin K.

APPLICANT: del Val, Greg

APPLICANT: Desjarlais, John R.

APPLICANT: Helfetz, Peter

APPLICANT: Muchhal, Umesh

TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity

FILE REFERENCE: A-71457-2/RFT/RMS/RMK

CURRENT APPLICATION NUMBER: US/10/141,531

PRIOR FILING DATE: 2002-05-06

PRIOR APPLICATION NUMBER: US 60/370,609

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US 60/289,029

PRIOR FILING DATE: 2001-05-04

NUMBER OF SEQ ID NOS: 208

SOFTWARE: Patentin version 3.1
SEQ ID NO 23
LENGTH: 135
TYPE: PRT
ORGANISM: Bacillus subtilis
US-10-141-531-23

Query Match 84.4%; Score 27; DB 6; Length 135;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6

Db 24 KEIHEE 29

RESULT 5
US-10-141-531-66

Sequence 66, Application US/10141531

GENERAL INFORMATION:

APPLICANT: Briggs, Steven P.

APPLICANT: Dalmia, Bipin K.

APPLICANT: del Val, Greg

APPLICANT: Desjarlais, John R.

APPLICANT: Helfetz, Peter

APPLICANT: Muchhal, Umesh

TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity

FILE REFERENCE: A-71457-2/RFT/RMS/RMK

CURRENT APPLICATION NUMBER: US/10/141,531

PRIOR FILING DATE: 2002-05-06

PRIOR APPLICATION NUMBER: US 60/370,609

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US 60/289,029

PRIOR FILING DATE: 2001-05-04

NUMBER OF SEQ ID NOS: 208

SOFTWARE: Patentin version 3.1

SEQ ID NO 66

LENGTH: 315

TYPE: PRT

ORGANISM: Bacillus subtilis

US-10-141-531-66

Query Match 84.4%; Score 27; DB 6; Length 315;

Best Local Similarity 66.7%; Pred. No. 68;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6

Db 204 KEIHEE 209

RESULT 6
US-10-141-531-85

Sequence 85, Application US/10141531

GENERAL INFORMATION:

APPLICANT: Briggs, Steven P.

APPLICANT: Dalmia, Bipin K.

APPLICANT: del Val, Greg

APPLICANT: Desjarlais, John R.

APPLICANT: Helfetz, Peter

APPLICANT: Muchhal, Umesh

TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity

FILE REFERENCE: A-71457-2/RFT/RMS/RMK

CURRENT APPLICATION NUMBER: US/10/141,531

PRIOR FILING DATE: 2002-05-06

PRIOR APPLICATION NUMBER: US 60/370,609

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US 60/289,029

PRIOR FILING DATE: 2001-05-04

NUMBER OF SEQ ID NOS: 208

SOFTWARE: Patentin version 3.1

;; SEQ ID NO 85
;; LENGTH: 316
;; TYPE: PRF
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-85

Query Match 84.4%; Score 27; DB 6; Length 316;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRYHEE 6
DB 205 KEIHEE 210

RESULT 7
US-10-141-531-94
;; Sequence 94, Application US/10141531
;; GENERAL INFORMATION:
;; APPLICANT: Briggs, Steven P.
;; APPLICANT: Dalma, Bipin K.
;; APPLICANT: del Val, Greg
;; APPLICANT: Desjarlais, John R.
;; APPLICANT: Helfetz, Peter
;; APPLICANT: Muchhal, Umesh
;; APPLICANT: Mughl, Peter
;; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
;; FILE REFERENCE: A-71457-2/RET/RMS/RMK
;; CURRENT APPLICATION NUMBER: US/10/141,531
;; CURRENT FILING DATE: 2002-05-06
;; PRIOR APPLICATION NUMBER: US 60/370,609
;; PRIOR FILING DATE: 2002-04-05
;; PRIOR APPLICATION NUMBER: US 60/289,029
;; PRIOR FILING DATE: 2001-05-04
;; NUMBER OF SEQ ID NOS: 208
;; SOFTWARE: Patent version 3.1
;; SEQ ID NO 94
;; LENGTH: 316
;; TYPE: PRF
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-94

Query Match 84.4%; Score 27; DB 6; Length 316;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRYHEE 6
DB 205 KEIHEE 210

RESULT 8
US-10-141-531-103
;; Sequence 103, Application US/10141531
;; GENERAL INFORMATION:
;; APPLICANT: Briggs, Steven P.
;; APPLICANT: Dalma, Bipin K.
;; APPLICANT: del Val, Greg
;; APPLICANT: Desjarlais, John R.
;; APPLICANT: Helfetz, Peter
;; APPLICANT: Muchhal, Umesh
;; APPLICANT: Mughl, Peter
;; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
;; FILE REFERENCE: A-71457-2/RET/RMS/RMK
;; CURRENT APPLICATION NUMBER: US/10/141,531
;; CURRENT FILING DATE: 2002-05-06
;; PRIOR APPLICATION NUMBER: US 60/370,609
;; PRIOR FILING DATE: 2002-04-05
;; PRIOR APPLICATION NUMBER: US 60/289,029

;; PRIOR FILING DATE: 2001-05-04
;; NUMBER OF SEQ ID NOS: 208
;; SOFTWARE: Patent version 3.1
;; SEQ ID NO 103
;; LENGTH: 316
;; TYPE: PRF
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-103

Query Match 84.4%; Score 27; DB 6; Length 316;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRYHEE 6
DB 205 KEIHEE 210

RESULT 9
US-10-141-531-112
;; Sequence 112, Application US/10141531
;; GENERAL INFORMATION:
;; APPLICANT: Briggs, Steven P.
;; APPLICANT: Dalma, Bipin K.
;; APPLICANT: del Val, Greg
;; APPLICANT: Desjarlais, John R.
;; APPLICANT: Helfetz, Peter
;; APPLICANT: Muchhal, Umesh
;; APPLICANT: Mughl, Peter
;; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
;; FILE REFERENCE: A-71457-2/RET/RMS/RMK
;; CURRENT APPLICATION NUMBER: US/10/141,531
;; CURRENT FILING DATE: 2002-05-06
;; PRIOR APPLICATION NUMBER: US 60/370,609
;; PRIOR FILING DATE: 2002-04-05
;; PRIOR APPLICATION NUMBER: US 60/289,029
;; PRIOR FILING DATE: 2001-05-04
;; NUMBER OF SEQ ID NOS: 208
;; SOFTWARE: Patent version 3.1
;; SEQ ID NO 112
;; LENGTH: 316
;; TYPE: PRF
;; ORGANISM: Artificial sequence
;; FEATURE:
;; OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-112

Query Match 84.4%; Score 27; DB 6; Length 316;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRYHEE 6
DB 205 KEIHEE 210

RESULT 10
US-10-141-531-121
;; Sequence 121, Application US/10141531
;; GENERAL INFORMATION:
;; APPLICANT: Briggs, Steven P.
;; APPLICANT: Dalma, Bipin K.
;; APPLICANT: del Val, Greg
;; APPLICANT: Desjarlais, John R.
;; APPLICANT: Helfetz, Peter
;; APPLICANT: Muchhal, Umesh
;; APPLICANT: Mughl, Peter
;; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
;; FILE REFERENCE: A-71457-2/RET/RMS/RMK
;; CURRENT APPLICATION NUMBER: US/10/141,531
;; CURRENT FILING DATE: 2002-05-06


```
; PRIOR APPLICATION NUMBER: US 60/370,609
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/289,029
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 121
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-121

Query Match      84.4%; Score 27; DB 6; Length 316;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
   1:|||||
Db 205 KEIHEE 210

RESULT 11
US-10-141-531-130
; Sequence 130, Application US/10141531
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Daimla, Bipin K.
; APPLICANT: del Val, Greg
; APPLICANT: Desjarlais, John R.
; APPLICANT: Helfetz, Peter
; APPLICANT: Luginduhl, Peter
; APPLICANT: Muchhal, Umesh
; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
; FILE REFERENCE: A-71457-2/RT/RMS/RMK
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US 60/370,609
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/289,029
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 130
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-130

Query Match      84.4%; Score 27; DB 6; Length 316;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
   1:|||||
Db 205 KEIHEE 210

RESULT 12
US-10-141-531-139
; Sequence 139, Application US/10141531
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Daimla, Bipin K.
; APPLICANT: del Val, Greg
; APPLICANT: Desjarlais, John R.
; APPLICANT: Helfetz, Peter
; APPLICANT: Luginduhl, Peter
; APPLICANT: Muchhal, Umesh
; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
```

```
; FILE REFERENCE: A-71457-2/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/141,531
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US 60/370,609
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/289,029
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 139
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-139

Query Match      84.4%; Score 27; DB 6; Length 316;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
   1:|||||
Db 205 KEIHEE 210

RESULT 13
US-10-141-531-148
; Sequence 148, Application US/10141531
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Daimla, Bipin K.
; APPLICANT: del Val, Greg
; APPLICANT: Desjarlais, John R.
; APPLICANT: Helfetz, Peter
; APPLICANT: Luginduhl, Peter
; APPLICANT: Muchhal, Umesh
; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
; FILE REFERENCE: A-71457-2/RT/RMS/RMK
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US/10/141,531
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US 60/370,609
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/289,029
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 148
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-148

Query Match      84.4%; Score 27; DB 6; Length 316;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
   1:|||||
Db 205 KEIHEE 210

RESULT 14
US-10-141-531-157
; Sequence 157, Application US/10141531
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Daimla, Bipin K.
; APPLICANT: del Val, Greg
; APPLICANT: Desjarlais, John R.
; APPLICANT: Helfetz, Peter
```

APPLICANT: Lugimbuhl, Peter
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
FILE REFERENCE: A-71457-2/RFT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/141,531
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: US 60/370,609
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: US 60/289,029
PRIOR FILING DATE: 2001-05-04
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PatentIn version 3.1
SEQ ID NO 157
LENGTH: 316
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-157

Query Match 84.4%; Score 27; DB 6; Length 316;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
I:::III
DB 205 KEIHEE 210

RESULT 15
US-10-141-531-166
Sequence 166, Application US/10141531
GENERAL INFORMATION:
APPLICANT: Briggs, Steven P.
APPLICANT: Dalmia, Bipin K.
APPLICANT: del Val, Greg
APPLICANT: Desjarlais, John R.
APPLICANT: Helletz, Peter
APPLICANT: Lugimbuhl, Peter
APPLICANT: Muchhal, Umesh
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
FILE REFERENCE: A-71457-2/RFT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/141,531
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: US 60/370,609
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: US 60/289,029
PRIOR FILING DATE: 2001-05-04
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PatentIn version 3.1
SEQ ID NO 166
LENGTH: 316
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-166

Query Match 84.4%; Score 27; DB 6; Length 316;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
I:::III
DB 205 KEIHEE 210

Search completed: November 13, 2002, 13:39:39
Job time : 2.42553 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using .sw model

Run on: November 13, 2002, 13:13:25 : Search time 9.31915 Seconds
(without alignments)
61.893 Million cell updates/sec

Title: US-09-856-086-5

Perfect score: 32

Sequence: 1 KKVHEE 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	68	2 A49830	ventricular myosin
2	32	100.0	144	2 E96618	hypothetical prote
3	32	100.0	287	2 A21762	neurofilament trip
4	32	100.0	298	1 DMHY	desmin - golden ha
5	32	100.0	308	2 D98095	conserved hypothet
6	32	100.0	463	1 DMCH	desmin - chicken
7	32	100.0	469	1 DMHU	desmin - human
8	32	100.0	469	2 T52469	desmin - rat
9	32	100.0	469	2 A24783	desmin - golden ha
10	32	100.0	469	2 A54104	desmin - mouse
11	32	100.0	543	1 OFMSL	neurofilament trip
12	32	100.0	544	2 S07144	neurofilament trip
13	32	100.0	544	2 B44841	low molecular weigh
14	32	100.0	548	1 OFPGL	neurofilament trip
15	32	100.0	554	1 JWO094	neurofilament trip
16	32	100.0	556	2 B46024	neurofilament-L su
17	31	96.9	329	2 F64356	translation initia
18	31	96.9	410	2 T19995	hypothetical prote
19	31	96.9	458	2 A43554	desmin - African c
20	31	96.9	807	2 T00990	hypothetical prote
21	31	96.9	1000	2 JEO110	mitotic control pr
22	31	96.9	3135	2 A48584	transmission block
23	29	90.6	216	2 H72277	hypothetical prote
24	29	90.6	251	2 T38625	urophosphorylgen
25	29	90.6	262	2 T44234	probable short-cha
26	29	90.6	265	2 T44234	hypothetical prote
27	29	90.6	270	2 T44049	hypothetical prote
28	29	90.6	333	2 F70341	sulfur oxygenase r
29	29	90.6	333	2 G84531	hypothetical prote

30	29	90.6	364	2 H83090	conserved hypothet
31	29	90.6	387	2 A86302	hypothetical prote
32	29	90.6	424	2 S14887	peripherin (clone
33	29	90.6	438	2 S07823	peripherin (clone
34	29	90.6	456	2 A60090	peripherin - Afric
35	29	90.6	458	2 S09228	intermediate filam
36	29	90.6	466	2 A25074	intermediate filam
37	29	90.6	468	2 JN0016	peripherin - human
38	29	90.6	469	2 A44841	low molecular weigh
39	29	90.6	471	2 A55185	intermediate filam
40	29	90.6	472	2 A51720	intermediate filam
41	29	90.6	509	2 T41034	probable WD-domain
42	29	90.6	665	2 S75291	cell division prot
43	29	90.6	767	2 S63220	probable membrane
44	29	90.6	823	2 B81282	probable integral
45	29	90.6	931	2 G69613	ATP-dependent DNA

ALIGNMENTS

RESULT 1
A49830 ventricular myosin heavy chain 1 - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
C:Accession: A49830
R:Blaesa, J.G.; Bader, D.
Dev. Biol. 148, 355-364, 1991
A>Title: Identification and characterization of a ventricular-specific avian myosin h
A:Reference number: A49830; MUID:92038448; PMID:1936571
A:Accession: A49830
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-68 <BIS>
A:Cross-references: GB:S64689; NID:9238622; PIDN:AAB20271.1; PID:9238623
A>Note: sequence extracted from NCBI backbone (NCBIN:64689, NCBI:P:64691)
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: ATP

Query Match 100.0%; Score 32; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
Db 62 KKVHEE 67

RESULT 2
E96618 hypothetical protein F9K23.1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E96618
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hultar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malt, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.H.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E96618
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-144 <STO>
A:Cross-references: GB:AE005173; NID:g11034967; PIDN:AGC27122.1; GSPDB:GN00141
C:Genetics:
A:Gene: F9K23.1

A:Map position: 1

Query Match 100.0%; Score 32; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKVHEE 6
|||||
Db 89 KKVHEE 94

RESULT 3

neurofilament triplet L protein - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)
C:Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 13-Aug-1999
A:Accession: A21762

R:Julien, J.P.; Ramachandran, K.; Grosveld, F.

Biochim. Biophys. Acta 825, 398-404, 1985

A:Title: Cloning of a cDNA encoding the smallest neurofilament protein from the rat.

A:Reference number: A21762; MUID:85252830; PMID:925999

A:Accession: A21762

A:Status: Preliminary

A:Molecule type: mRNA

A:Cross-references: GB:M25638; NID:g205683; PID:AAA1694.1; PID:g205684

C:Superfamily: Cytoskeletal keratin

C:Keywords: coiled coil

Query Match 100.0%; Score 32; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKVHEE 6
|||||
Db 27 KKVHEE 32

RESULT 4

desmin - golden hamster (fragment)

C:Species: Mesocricetus auratus (golden hamster)
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 22-Jun-1999
A:Accession: A02956

R:Quax, W.; van den Heuvel, R.; Egberts, W.V.; Quax-Deuker, Y.; Bloemendal, H.

Proc. Natl. Acad. Sci. U.S.A. 81, 5870-5874, 1984

A:Title: Intermediate filament cDNAs from BHK-21 cells: demonstration of distinct genes

A:Reference number: A02956; MUID:85014890; PMID:6091127

A:Accession: A02956

A:Molecule type: mRNA

A:Residues: 1-298 <QUN>

A:Cross-references: GB:K02407; NID:g191356; PID:AAA7071.1; PID:g387070

A:Experimental source: baby hamster kidney cells, BHK-21

C:Comment: There is a single gene for desmin in the hamster genome.

C:Superfamily: Cytoskeletal keratin

C:Keywords: coiled coil; intermediate filament

F:245-298/Domain: tail <TLE>

Query Match 100.0%; Score 32; DB 1; Length 298;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKVHEE 6
|||||
Db 68 KKVHEE 73

RESULT 5

conserved hypothetical protein spr1790 [imported] - Streptococcus pneumoniae (strain R6)

C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 31-Dec-2001

C:Accession: D98095

R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Deroff, B.;
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren,
Y.P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Balte, R.H.; Jaskunas,

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: D98095

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-308 <KUR>

A:Cross-references: GB:AE07317; PID:AA100593.1; PID:g15459474; GSPDB:GN00174

C:Genetics:

A:Gene: spr1790

C:Superfamily: stage III sporulation protein; stage III sporulation protein homolog.

Query Match 100.0%; Score 32; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKVHEE 6
|||||
Db 258 KKVHEE 263

RESULT 6

DMCH

desmin - chicken

N:Alternate names: type III intermediate filament

C:Species: Gallus gallus (chicken)

C:Date: 18-Dec-1981 #sequence_revision 12-Apr-1996 #text_change 10-Dec-1999

A:Accession: A90969; A94014; JCI459; S02448; A32858; S23189; A02957

R:Geisler, N.; Weber, K.

EMBO J. 1, 1649-1656, 1982

A:Title: The amino acid sequence of chicken muscle desmin provides a common structure

A:Reference number: A90969; MUID:84207925; PMID:6202512

A:Accession: A90969

A:Molecule type: protein

A:Residues: 1-463 <GEI>

R:Capetanaki, Y.G.; Ngai, J.; Lazarides, E.

Proc. Natl. Acad. Sci. U.S.A. 81, 6909-6913, 1984

A:Title: Characterization and regulation in the expression of a gene coding for the

A:Reference number: A94014; MUID:85063701; PMID:6594672

A:Accession: A94014

A:Molecule type: mRNA

A:Residues: 365-463 <CAP>

A:Cross-references: GB:K02445; NID:g211727; PID:AAA48751.1; PID:g211728

R:Kusubata, M.; Matsuo, Y.; Tsujimura, K.; Ito, H.; Ando, S.; Kamijo, M.; Yasuda,

Biochem. Biophys. Res. Commun. 190, 927-934, 1993

A:Title: cdc2 Kinase phosphorylation of desmin at three serine/threonine residues 11

A:Reference number: JCI459; MUID:93176201; PMID:8439342

A:Accession: JCI459

A:Molecule type: protein

A:Residues: 5-9,16-27,63-67 <KUS>

R:Geisler, N.; Weber, K.

EMBO J. 7, 15-20, 1988

A:Title: Phosphorylation of desmin in vitro inhibits formation of intermediate filaments

A:Reference number: S02448; MUID:88196075; PMID:3359992

A:Accession: S02448

A:Molecule type: protein

A:Residues: 1-69 <GE2>

R:Kittamura, S.; Ando, S.; Shibata, M.; Tanabe, K.; Sato, C.; Inagaki, M.

J. Biol. Chem. 264, 5674-5678, 1989

A:Title: Protein kinase C phosphorylation of desmin at four serine residues within 1

A:Reference number: A32858; MUID:89174618; PMID:2494168

A:Accession: A32858

A:Molecule type: protein

A:Residues: 10-14;28-42;49-59 <KIT>

R:Geisler, N.; Schuenemann, J.; Weber, K.

Eur. J. Biochem. 206, 841-852, 1992

A:Title: Chemical cross-linking indicates a staggered and antiparallel protofilament

A:Reference number: S23189; MUID:92299013; PMID:1606966

A:Accession: S23189

A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 110-118;255-266;274-282;393-401 <GE3>
 C:Comment: This protein was isolated from chicken gizzard.
 C:Comment: There appears to be a single desmin gene in the haploid chicken genome that is predominantly in skeletal, cardiac, and most types of smooth muscle cells during myogenic
 C:Comment: Desmin intermediate filaments are found in the cytoplasm of cultured myogenic
 C:Comment: myofibrils to each other and to the plasma membrane from the periphery of the Z-line
 C:Comment: The molecule contains three structurally distinct domains. The surface-exposed
 C:Superfamily: cytoskeletal keratin
 C:Keywords: blocked amino end; coiled coil; intermediate filament; muscle; phosphoprotein
 F:1-99/Domain: head <HED>
 F:100-407/Domain: rod <ROD>
 F:100-132/Region: coil 1A
 F:133-143/Region: linker 1
 F:143-243/Region: coil 1B
 F:244-259/Region: linker 12
 F:260-278/Region: coil 2A
 F:279-286/Region: linker 2
 F:287-407/Region: coil 2B
 F:345/Region: stuffer
 F:408-463/Domain: tail <END>
 F:1/Modified site: blocked amino end (Ser) (probably acetylated) #status experimental
 F:6,42/Binding site: phosphate (Ser) (covalent) (by cd2 kinase) #status experimental
 F:12,29,38,56/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status exp
 F:29,35,50/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status p
 F:64/Binding site: phosphate (Thr) (covalent) (by cd2 kinase) #status experimental

Query Match 100.0%; Score 32; DB 1; Length 463;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 DB 231 KKVHEE 236

RESULT 7

DNHU
 desmin - human
 M:Alternate names: type III intermediate filament
 C:Species: Homo sapiens (man)
 C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 10-Dec-1999
 C:Accession: J00063
 R:1, 2; Lilienbaum, A.; Butler-Browne, G.; Paulin, D.
 Gene 78, 243-254, 1989
 A:Title: Human desmin coding gene: complete nucleotide sequence, characterization and re
 A:Reference number: J00063; MUID:89378751; PMID:2673923
 A:Accession: J00063
 A:Molecule type: DNA
 A:Residues: 1-469 <L12>
 A:Cross-references: GB:M63391; GB:M26935; GB:M58168; GB:M59379; GB:M65071; GB:X53154; NT
 A>Note: the introns of this gene contain 1.2K of repetitive sequences belonging to the h
 C:Comment: Desmin intermediate filaments are found in the cytoplasm of cultured myogenic
 C:Comment: myofibrils to each other and to the plasma membrane from the periphery of the Z-line
 C:Comment: The molecule contains three structurally distinct domains. The surface-exposed
 C:Superfamily: cytoskeletal keratin
 C:Keywords: blocked amino end; coiled coil; intermediate filament; muscle; phosphoprotein
 F:1-108/Domain: head <HED>
 F:109-415/Domain: rod <ROD>
 F:416-469/Domain: tail <END>
 F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
 F:13,48/Binding site: phosphate (Ser) (covalent) (by cd2 kinase) #status predicted
 F:29,35,50/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 F:45,60/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status pred

Query Match 100.0%; Score 32; DB 1; Length 463;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 DB 239 KKVHEE 244

RESULT 8

desmin - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 13-Aug-1999
 C:Accession: I52469
 R:van Groningen, J.J.; Bloemers, H.P.; Swart, G.W.
 Biochim. Biophys. Acta 1217, 107-109, 1994
 A:Title: Rat desmin gene structure and expression.
 A:Reference number: I52469; MUID:94114566; PMID:8286410
 A:Accession: I52469
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-469 <RES>
 A:Cross-references: EMBL:X73524; NID:9452778; PIDN:CA51920.1; PID:9452779
 C:Superfamily: cytoskeletal keratin

Query Match 100.0%; Score 32; DB 2; Length 469;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 DB 239 KKVHEE 244

RESULT 9

A24783
 desmin - golden hamster
 C:Species: Mesocricetus auratus (golden hamster)
 C:Date: 22-Jul-1987 #sequence_revision 22-Jul-1987 #text_change 13-Aug-1999
 C:Accession: A24783
 R:Quax, W.; van den Broek, L.; Egberts, W.V.; Ramaekers, F.; Bloemendaal, H.
 Cell 43, 327-338, 1985
 A:Title: Characterization of the hamster desmin gene: expression and formation of des
 A:Reference number: A24783; MUID:86079506; PMID:3855248
 A:Accession: A24783
 A:Molecule type: DNA
 A:Residues: 1-469 <QUA>
 A:Cross-references: GB:M12104; NID:9191360; PIDN:AAA37072.1; PID:9387071
 C:Genetics:
 A:Introns: 192/2; 212/3; 244/3; 298/3; 340/3; 414/2; 429/1; 456/3
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil

Query Match 100.0%; Score 32; DB 2; Length 469;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 DB 239 KKVHEE 244

RESULT 10

A54104
 desmin - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 19-Apr-1996
 C:Accession: A54104; S31404
 R:Li, H.; Choudhary, S.K.; Milner, D.J.; Munir, M.I.; Kusk, I.R.; Capetanaki, Y.
 J. Cell Biol. 124, 827-841, 1994
 A:Title: Inhibition of desmin expression blocks myoblast fusion and interferes with t

A:Reference number: A54104; MUID:94165148; PMID:8120103
 A:Accession: A54104
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-469 <LIN>
 A:Cross-references: GB:122550
 R:Li, H.; Capetanakl, Y.
 Submitted to the EMBL Data Library, December 1992
 A:Description: Regulation of the Mouse Desmin Gene: Transactivation by MyoD, Myogenin, M
 A:Reference number: S31404
 A:Accession: S31404
 A:Molecule type: DNA
 A:Residues: 1-40 <LIN>
 A:Cross-references: EMBL:216892
 C:Superfamily: cytoskeletal keratin
 C:Keywords: muscle

Query Match 100.0%; Score 32; DB 2; Length 469;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 DB 239 KKVHEE 244

RESULT 11
 OFMSL
 neurofilament triplet L protein - mouse
 N:Alternate names: 68k neurofilament protein; NF-L(low) protein; type IV IF protein
 C:Species: Mus musculus (house mouse)
 C>Date: 31-Mar-1988 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999
 C:Accession: A25227; A26562; A43772; A41012; I55316
 R:Lewis, S.A.; Cowan, N.J.
 M1. Cell. Biol. 6, 1529-1534, 1986
 A:Title: Anomalous placement of introns in a member of the intermediate filament multigene
 A:Reference number: A25227; MUID:87064433; PMID:3785173
 A:Accession: A25227
 A:Molecule type: DNA
 A:Residues: 1-343 <LEM>
 A:Cross-references: GB:M13016; NID:9200023; PIDN:AAA39810.1; PID:9387492
 A:Note: the authors translated the codon GGC for residue 5 as Ala, ACG for residue 88 as
 1 as Glu
 R:Lewis, S.A.; Cowan, N.J.
 J. Cell Biol. 100, 843-850, 1985
 A:Title: Genetics, evolution, and expression of the 68,000-mol-wt neurofilament protein:
 A:Reference number: A26562; MUID:8513134; PMID:3919033
 A:Accession: A26562
 A:Molecule type: mRNA
 A:Residues: 242-543 <LE2>
 A:Cross-references: GB:X02165
 A:Experimental source: Brain
 R:Julien, J.P.; Meyer, D.; Flavell, D.; Hurst, J.; Grosfeld, F.
 Brain Res. Mol. Brain Res. 1, 243-250, 1986
 A:Title: Cloning and developmental expression of the murine neurofilament gene family.
 A:Reference number: A43772
 A:Accession: A43772
 A:Molecule type: mRNA
 A:Residues: 1-5, 'Y', '7-8, 'Y', '10-64, 'W', '66-72, 'L', '74-98, 'D', '100-194, 'R', '196-202, '204-239, 'Y
 A:Cross-references: GB:M20480; NID:9200037; PIDN:AAA39814.1; PID:9200038
 A:Note: the authors translated the codon CCG for residue 195 as Ala
 R:Shing, R.K.; Nixon, R.A.
 J. Biol. Chem. 266, 18861-18867, 1991
 A:Title: Identification of Ser-55 as a major protein kinase A phosphorylation site on th
 A:Reference number: A41012; MUID:92011653; PMID:1717455
 A:Accession: A41012
 A:Molecule type: protein
 A:Residues: 52-57 <SIH>
 R:Nakahira, K.; Ikenaka, K.; Wada, K.; Tamura, T.
 J. Biol. Chem. 265, 19786-19791, 1990
 A:Title: Structure of the 68-kDa neurofilament gene and regulation of its expression.
 A:Reference number: I55316; MUID:91060592; PMID:2246261
 A:Accession: I55316

A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-5, 'Y', '7-8, 'Y', '10-28 <RES>
 A:Cross-references: GB:M55423; NID:9200027; PIDN:AAA39812.1; PID:9554245
 C:Comment: This is the most abundant of the three neurofilament proteins and, as th
 C:Genetics:
 A:introns: 349/3; 391/2; 498/1
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 P:2-72/Domain: head <RED>
 F:94-125/Domain: coil 1a, alpha-helical rod #status predicted <RIA>
 F:126-138/Region: linker 1
 F:139-234/Domain: coil 1b, alpha-helical rod #status predicted <R1B>
 F:235-256/Region: linker 12
 F:257-272/Domain: coil 2a, alpha-helical rod #status predicted <R2A>
 F:273-281/Region: linker 2
 F:282-401/Domain: coil 2b, alpha-helical rod #status predicted <R2B>
 F:404-543/Domain: tail <TAI>
 F:404-444/Region: tail subdomain a
 F:445-543/Region: tail subdomain b

Query Match 100.0%; Score 32; DB 1; Length 543;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 DB 224 KKVHEE 229

RESULT 12
 S07144
 neurofilament triplet L protein - human
 N:Alternate names: neurofilament light polypeptide (68k)
 N:Contains: Glu-50 brain peptide
 C:Species: Homo sapiens (man)
 C>Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 21-Jul-2000
 C:Accession: S07144; I52832; A60703
 R:Julien, J.P.; Grosfeld, F.; Yazdankhsh, K.; Flavell, D.; Meijer, D.; Mushynski, I
 Biochim. Biophys. Acta 909, 10-20, 1987
 A:Title: The structure of a human neurofilament gene (NF-L): a unique exon-intron o
 A:Reference number: S07144; MUID:87214213; PMID:3034332
 A:Accession: S07144
 A:Molecule type: DNA
 A:Residues: 1-544 <JUI>
 A:Cross-references: EMBL:X05608; NID:91495072; PIDN:CAA29097.1; PID:91279504
 A:Note: the authors translated the codon ARG for residue 366 as Asn
 R:Pospelova, V.A.; Pospelova, T.V.; Julien, J.P.
 Cell Growth Differ. 5, 187-196, 1994
 A:Title: AP-1 and Krox-24 transcription factors activate the neurofilament light ge
 A:Reference number: I52832; MUID:94235564; PMID:8180132
 A:Accession: I52832
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-10 <POS>
 A:Cross-references: GB:S70309; NID:9547176; PIDN:AAD14057.1; PID:94261757
 R:Nomata, Y.; Watanabe, T.; Wada, H.
 J. Biochem. 93, 825-831, 1983
 A:Title: Highly acidic proteins from human brain: purification and properties of Gl
 A:Reference number: A60703; MUID:83265667; PMID:6135695
 A:Accession: A60703
 A:Molecule type: protein
 A:Residues: 469-472, 'D', '474 <NOM>
 A:Experimental source: Glu-50 brain peptide
 A:Note: this acidic protein is named for its greater than fifty per cent glutamic a
 C:Genetics:
 A:Gene: GDB:NEFL, NFL
 A:Cross-references: GDB:120227; OMIM:162280
 A:Map position: 8p21-8p21
 A:introns: 349/3; 391/2; 498/1
 C:superfamily: cytoskeletal keratin
 C:Keywords: brain; coiled coil; intermediate filament
 F:469-544/Product: Glu-50 peptide #status predicted <E50>

Query Match 100.0%; Score 32; DB 2; Length 544;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
 |||||
 Db 225 KKVHEE 230

RESULT 13

B44841 low molecular weight neurofilament protein XNF-L - African clawed frog

C:Species: Xenopus laevis (African clawed frog)
 C>Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
 C:Accession: B44841
 R:Charnas, L.R.; Szaro, B.G.; Gainer, H.
 J. Neurosci. 12, 3010-3024, 1992
 A>Title: Identification and developmental expression of a novel low molecular weight neurofilament protein from the African clawed frog
 A:Reference number: A44841; MUID:93356194; PMID:1494944
 A:Accession: B44841
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-544 <CHA>
 A:Experimental source: brain
 A>Note: sequence inconsistent with the nucleotide translation
 A>Note: sequence extracted from NCBI backbone (NCBIN:110225, NCBI:110226)
 C:Superfamily: cytoskeletal keratin

Query Match 100.0%; Score 32; DB 2; Length 544;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
 |||||
 Db 218 KKVHEE 223

RESULT 14

OPFGL

neurofilament triplet L protein - pig

N:Alternate names: 68K neurofilament protein
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 15-Nov-1984 #sequence_revision 28-May-1986 #text_change 10-Oct-1997
 C:Accession: A91337; A90973; A34569; A02963
 R:Geisler, N.; Plesmann, U.; Weber, K.
 FEBS Lett. 182, 475-478, 1985
 A>Title: The complete amino acid sequence of the major mammalian neurofilament protein
 A:Reference number: A91337; MUID:85154583; PMID:3920075
 A:Accession: A91337
 A:Molecule type: protein
 A:Residues: 1-547 <GE1>
 R:Geisler, N.; Kaufmann, E.; Fischer, S.; Plesmann, U.; Weber, K.
 EMBO J. 2, 1295-1302, 1983
 A>Title: Neurofilament architecture combines structural principles of intermediate filament
 A:Reference number: A90973
 A:Accession: A90973
 A:Molecule type: protein
 A:Residues: 1-82;278-548 <GE2>
 A>Note: residue 322 is either lysine or arginine
 R:Gonda, Y.; Nishizawa, K.; Ando, S.; Kitamura, S.; Minoura, Y.; Nish, Y.; Inagaki, M.
 Biochem. Biophys. Res. Commun. 167, 1316-1325, 1990
 A>Title: Involvement of protein kinase C in the regulation of assembly-disassembly of neurofilament
 A:Reference number: A34569; MUID:90211318; PMID:2108674
 A:Accession: A34569
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 9-14;23-29;30-53 <GON>

C:Comment: Mammalian neurofilaments usually contain three polypeptides, L, M, and H (with the H being the heaviest). The neurofilament protein is a conserved alpha-helical region, whose helix is flanked by two long flexible regions. The amino-terminal headpiece is basic with a high content of hydroxyamino acids and beta turns; domain b is acidic and rich in glutamic acid and lysine residues.
 C:Comment: The extra mass and high charge density that distinguish the neurofilament protein

charged scaffolding structure suitable for interaction with other neuronal component
 C:Comment: The boundaries of the domains between residues 70-92 and 399-402 are not yet defined.

C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:1-70/Domain: head <HED>
 F:92-123/Domain: coil 1a, alpha-helical rod <R1A>
 F:137-232/Domain: coil 1b, alpha-helical rod <R1B>
 F:255-399/Domain: coil 2, alpha-helical rod <R2>
 F:402-548/Domain: tail <TAI>
 F:402-442/Region: tail subdomain a
 F:443-548/Region: tail subdomain b

Query Match 100.0%; Score 32; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
 |||||
 Db 222 KKVHEE 227

RESULT 15

JM0094 neurofilament protein-L - bovine

N:Alternate names: NF-L
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 07-May-1999
 C:Accession: JM0094
 R:Hashimoto, R.; Nakamura, Y.; Goto, H.; Wada, Y.; Sakoda, S.; Kalbuehl, K.; Inagaki, M.
 Biochem. Biophys. Res. Commun. 245, 407-411, 1998
 A>Title: Domain- and site-specific phosphorylation of bovine NF-L by Rho-associated kinase
 A:Reference number: JM0094; MUID:98238650; PMID:9571164
 A:Accession: JM0094
 A:Molecule type: protein
 A:Residues: 1-554 <HAS>
 C:Comment: Domain- and site-specific phosphorylation by Rho-kinase regulates the assembly of neurofilament
 C:Superfamily: cytoskeletal keratin

Query Match 100.0%; Score 32; DB 2; Length 554;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHEE 6
 |||||
 Db 223 KKVHEE 228

Search completed: November 13, 2002, 13:23:28
 Job time : 9.46201 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 4.97872 Seconds
(without alignments)
49.984 Million cell updates/sec

Title: US-09-856-086-5

Perfect score: 32

Sequence: 1 KKVHEE 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	463	1	DESM_CHICK
2	32	100.0	468	1	DESM_MESNU
3	32	100.0	468	1	DESM_MOUSE
4	32	100.0	468	1	DESM_MOUSE
5	32	100.0	469	1	DESM_BOVIN
6	32	100.0	469	1	DESM_HUMAN
7	32	100.0	470	1	DESM_PIG
8	32	100.0	541	1	NFL_RAT
9	32	100.0	542	1	NFL_MOUSE
10	32	100.0	543	1	NFL_HUMAN
11	32	100.0	544	1	NFL_XENLA
12	32	100.0	548	1	NFL_PIG
13	32	100.0	554	1	NFL_BOVIN
14	32	100.0	555	1	NFL_CORTA
15	32	100.0	555	1	NFL_MOUSE
16	32	100.0	555	1	NFL_MOUSE
17	32	100.0	555	1	NFL_MOUSE
18	32	100.0	555	1	NFL_MOUSE
19	32	100.0	555	1	NFL_MOUSE
20	32	100.0	555	1	NFL_MOUSE
21	32	100.0	555	1	NFL_MOUSE
22	32	100.0	555	1	NFL_MOUSE
23	32	100.0	555	1	NFL_MOUSE
24	32	100.0	555	1	NFL_MOUSE
25	32	100.0	555	1	NFL_MOUSE
26	32	100.0	555	1	NFL_MOUSE
27	32	100.0	555	1	NFL_MOUSE
28	32	100.0	555	1	NFL_MOUSE
29	32	100.0	555	1	NFL_MOUSE
30	32	100.0	555	1	NFL_MOUSE
31	32	100.0	555	1	NFL_MOUSE
32	32	100.0	555	1	NFL_MOUSE
33	32	100.0	555	1	NFL_MOUSE

34	28	87.5	357	1	RECT_LACLC	P50925 lactococcus
35	28	87.5	358	1	RECT_LACLC	O9CE70 lactococcus
36	28	87.5	428	1	GFAP_BOVIN	O28115 bos taurus
37	28	87.5	430	1	GFAP_RAT	P47819 rattus norv
38	28	87.5	432	1	GFAP_HUMAN	P14136 homo sapien
39	28	87.5	493	1	YDAK_CAEEL	P90771 caenorhabdi
40	28	87.5	513	1	CP12_MOUSE	P00186 mus musculu
41	28	87.5	513	1	CP12_MOUSE	P04799 rattus norv
42	28	87.5	538	1	RO60_MOUSE	O08848 mus musculu
43	28	87.5	553	1	C301_DROME	O9V666 drosophila
44	28	87.5	978	1	RA50_AQUAE	O67124 aquifex aeo
45	28	87.5	1115	1	IRE1_YEAST	P32361 saccharomyc

ALIGNMENTS

RESULT 1	ID	DESM_CHICK	STANDARD:	PRT:	463 AA.
AC	P02542;				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	23-OCT-1986 (Rel. 02, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Desmin.				
GN	DES.				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
OC	Gallus.				
OX	NCBI_TaxID-9031;				
RN	[1]				
RP	SEQUENCE.				
RC	Tissue=clzard;				
RX	MEDLINE-84207925; PubMed-6202512;				
RA	Gelsler N., Weber K.;				
RT	"The amino acid sequence of chicken muscle desmin provides a common				
RT	structural model for intermediate filament proteins.";				
RL	EMBO J. 1:1649-1656(1982).				
RN	[2]				
RP	SEQUENCE OF 1-88 AND 254-415.				
RX	MEDLINE-83025086; PubMed-6889923;				
RA	Gelsler N., Kaufmann E., Weber K.;				
RT	"Proteinchemical characterization of three structurally distinct				
RT	domains along the protofilament unit of desmin 10 nm filaments.";				
RL	Cell 30:277-286(1982).				
RN	[3]				
RP	SEQUENCE OF 324-463.				
RX	MEDLINE-82037793; PubMed-6945574;				
RA	Gelsler N., Weber K.;				
RT	"Comparison of the proteins of two immunologically distinct				
RT	intermediate-sized filaments by amino acid sequence analysis: desmin				
RL	and vimentin.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 78:4120-4123(1981).				
RN	[4]				
RP	SEQUENCE OF 366-463 FROM N.A.				
RX	MEDLINE-85063701; PubMed-6594672;				
RA	Capetanaki Y.G., Ngai J., Lazarides E.;				
RT	"Characterization and regulation in the expression of a gene coding				
RT	for the intermediate filament protein desmin.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 81:6909-6913(1984).				
CC	-1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN				
CC	MUSCLE CELLS. IN ADULT STRATIFIED MUSCLE THEY FORM A FIBROUS				
CC	NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA				
CC	MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.				
CC	-1- SUBUNIT: HOMOPOLYMER.				
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic.				
CC	-1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				

modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: K02445; AAA48751.1; -
DR PIR: A02957; DMCH.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Muscle protein.
FT MOD_RES 1 1 BLOCKED.
FT DOMAIN 1 99 HEAD.
FT DOMAIN 100 403 ROD.
FT DOMAIN 404 463 TAIL.
FT DOMAIN 100 132 COIL 1A.
FT DOMAIN 133 142 COIL 1B.
FT DOMAIN 143 243 LINKER 1.
FT DOMAIN 244 259 LINKER 12.
FT DOMAIN 260 278 COIL 2A.
FT DOMAIN 279 286 LINKER 2.
FT DOMAIN 287 403 COIL 2B.
FT SITE 345 345 STUTTER.
SO SEQUENCE 463 AA; 53279 MW; F9AFC2CF9CD111C CRC64;

Query Match 100.0%; Score 32; DB 1; Length 463;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
Db 231 KKVHEE 236

RESULT 2
DESM_MESAU STANDARD: PRT: 468 AA.
ID DESM_MESAU
AC P02541;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desmin.
GN DES.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_Taxid=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86079506; PubMed=3855248;
RA Quax W.J., van den Broek L., Egberts W.V., Rameekers F.,
Bloemendal H.;
RT "Characterization of the hamster desmin gene: expression and
formation of desmin filaments in nonmuscle cells after gene
transfer.";
RT Cell 43:327-338(1985).
RN [2]
RP SEQUENCE OF 171-468 FROM N.A.
RX MEDLINE=85014890; PubMed=6091127;
RA Quax W.J., van den Heuvel R., Egberts W.V., Quax-Jeukens Y.E.F.M.,
Bloemendal H.;
RT "Intermediate filament cDNAs from BHK-21 cells: demonstration of
distinct genes for desmin and vimentin in all vertebrate classes.";
RT Proc. Natl. Acad. Sci. U.S.A. 81:5970-5974(1984).
CC -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
CC -1- SUBCELLULAR LOCATION: CYTOSOL.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL: K02407; AAA37071.1; -
DR EMBL: M12104; AAA37072.1; JOINED.
DR EMBL: M12102; AAA37072.1; JOINED.
DR EMBL: M12103; AAA37072.1; JOINED.
DR PIR: A02956; DMHY.
DR PIR: A24783; A24783.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Muscle protein.
FT INIT_MET 0 0
FT DOMAIN 1 106 HEAD.
FT DOMAIN 107 410 ROD.
FT DOMAIN 411 468 TAIL.
FT DOMAIN 107 139 COIL 1A.
FT DOMAIN 140 149 LINKER 1.
FT DOMAIN 150 250 COIL 1B.
FT DOMAIN 251 266 LINKER 12.
FT DOMAIN 267 285 COIL 2A.
FT DOMAIN 286 293 LINKER 2.
FT DOMAIN 294 410 COIL 2B.
FT SITE 352 352 STUTTER.
SO SEQUENCE 468 AA; 53307 MW; 39159431C5908FB9 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 468;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
Db 238 KKVHEE 243

RESULT 3
DESM_MOUSE STANDARD: PRT: 468 AA.
ID DESM_MOUSE
AC P31001;
DT 01-JUL-1993 (Rel. 26, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desmin.
GN DES.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94165148; PubMed=8120103;
RA Li H., Choudhary S.K., Milner D.J., Munir M.I., Kulsk I.R.,
Capetanaki Y.;
RT "Inhibition of desmin expression blocks myoblast fusion and
interferes with the myogenic regulators MyoD and myogenin.";
RT J. Cell Biol. 124:827-841(1994).
RN [2]
RP SEQUENCE OF 1-39 FROM N.A.
RX STRAIN-BALB/c; TISSUE-Spleen;
RX MEDLINE=93181210; PubMed=8382796;
RA Li H., Capetanaki Y.;
RT "Regulation of the mouse desmin gene: transactivated by MyoD,
myogenin, MRF4 and Myf5.";
RT Nucleic Acids Res. 21:335-343(1993).
CC -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
CC -1- SUBUNIT: HOMOPOLYMER.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L22550; -; NOT_ANNOTATED_CDS.
 CC EMBL: 218892; CAA79330.1; -.
 CC PIR: S31404; S31404.
 CC SWISS-2DPAGE: P31001; MOUSE.
 CC MGD: MGI:94885; Des.
 CC InterPro: IPR001664; IF.
 CC Pfam: PF00038; Filament; 1.
 CC PROSITE: PS00226; IF; 1.
 CC Intermediate filament; Coiled coil; Muscle protein.
 CC KW INIT MET 0 0
 CC FT DOMAIN 1 107 HEAD.
 CC FT DOMAIN 108 410 ROD.
 CC FT DOMAIN 411 458 TAIL.
 CC FT DOMAIN 108 139 COIL 1A.
 CC FT DOMAIN 140 149 LINKER 1.
 CC FT DOMAIN 150 250 COIL 1B.
 CC FT DOMAIN 251 266 LINKER 12.
 CC FT DOMAIN 267 285 COIL 2A.
 CC FT DOMAIN 286 293 LINKER 2.
 CC FT DOMAIN 294 410 COIL 2B.
 CC FT SITE 352 352 STUTTER.
 CC FT SITE 44 47 POLY-SER.
 CC FT DOMAIN 44 47
 CC SQ SEQUENCE 468 AA; 53366 MW; FD276059E7E3BFB CRC64;
 CC
 CC Query Match 100.0%; Score 32; DB 1; Length 468;
 CC Best Local Similarity 100.0%; Pred. NO. 14;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Oy 1 KRYHEE 6
 CC Db 238 KRYHEE 243
 CC
 CC RESULT 4
 CC DESM_BOVIN STANDARD; PRT; 468 AA.
 CC ID DESM_BOVIN
 CC AC P48675;
 CC DT 01-FEB-1996 (Rel. 33; Created)
 CC DT 01-FEB-1996 (Rel. 33; Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40; Last annotation update)
 CC DE Desmln.
 CC OS Rattus norvegicus (Rat).
 CC CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC OX NCBI_TaxID=10116;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Histar; TISSUE=Arteria;
 CC RX MEDLINE=9411456; PubMed=8286410;
 CC RA van Groningen J.J.M., Bloemers H.P.J., Swart G.W.M.;
 CC RT "Rat desmln gene structure and expression";
 CC RL Blochin. Biophys. Acta 1217:107-109(1994).
 CC CC -1- FUNCTION: DESMLN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
 CC MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
 CC NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
 CC MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
 CC CC -1- SUBUNIT: HOMOPOLYMER.
 CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB011675; BAA25135.1; -.
 CC EMBL: AB011673; BAA25133.1; -.
 CC InterPro: IPR001664; IF.
 CC

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X73524; CAA51920.1; -.
 CC InterPro: IPR001664; IF.
 CC Pfam: PF00038; Filament; 1.
 CC PROSITE: PS00226; IF; 1.
 CC Intermediate filament; Coiled coil; Muscle protein.
 CC KW INIT MET 0 0
 CC FT DOMAIN 1 107 HEAD.
 CC FT DOMAIN 108 410 ROD.
 CC FT DOMAIN 411 458 TAIL.
 CC FT DOMAIN 108 139 COIL 1A.
 CC FT DOMAIN 140 149 LINKER 1.
 CC FT DOMAIN 150 250 COIL 1B.
 CC FT DOMAIN 251 266 LINKER 12.
 CC FT DOMAIN 267 285 COIL 2A.
 CC FT DOMAIN 286 293 LINKER 2.
 CC FT DOMAIN 294 410 COIL 2B.
 CC FT DOMAIN 410 410
 CC SQ SEQUENCE 468 AA; 53325 MW; EBD0667A5A7FA8B CRC64;
 CC
 CC Query Match 100.0%; Score 32; DB 1; Length 468;
 CC Best Local Similarity 100.0%; Pred. NO. 14;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Oy 1 KRYHEE 6
 CC Db 238 KRYHEE 243
 CC
 CC RESULT 5
 CC DESM_BOVIN STANDARD; PRT; 469 AA.
 CC ID DESM_BOVIN
 CC AC 062654; 062655;
 CC DT 15-DEC-1998 (Rel. 37; Created)
 CC DT 15-JUL-1999 (Rel. 38; Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40; Last annotation update)
 CC DE Desmln.
 CC OS Bos taurus (Bovine).
 CC CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC OX Bovidae; Bovinae; Bos.
 CC OX NCBI_TaxID=9913;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=Holstein; TISSUE=Muscle;
 CC RA Chikuni K., Tanabe R., Muroya S.;
 CC RT "Desmln structure as related to meat tenderness";
 CC RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC CC -1- FUNCTION: DESMLN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
 CC MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
 CC NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
 CC MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
 CC CC -1- SUBUNIT: HOMOPOLYMER.
 CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AB011675; BAA25135.1; -.
 CC EMBL: AB011673; BAA25133.1; -.
 CC InterPro: IPR001664; IF.
 CC


```

FT 22 24 GFP -> VRS (IN REF. 1 AND 2).
FT 38 38 G -> P (IN REF. 1 AND 2).
FT 118 122 FANYI -> SPIYM (IN REF. 1 AND 2).
FT 134 134 MISSING (IN REF. 1, 2 AND 3).
SQ SEQUENCE 469 AA; 53404 MW; 6A38116859A091B8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 469;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHHE 6
DB 239 KKVHHE 244

RESULT 7
DESM_PIG STANDARD; PRT; 470 AA.
AC P02340; 062656;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Desmin.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Chikuni K., Tanabe R., Muroya S.;
RT "Desmin structure as related to meat tenderness.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Tugue C.K., Sanchez-Serrano I., Smith B., Marklund L., Ernst C.;
RT "Muscle Esrs II: Cloning, sequencing and mapping the pig gene for the
intermediate filament protein desmin (DES).";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Longissimus muscle;
RA Beuzen N.D., Hall A.D., Gallagher A., Chang K.-C.;
RT "A polymorphic CT-repeat at the porcine desmin locus with an effect on
meat quality.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 333-470.
RA MEDLINE=82037793; PubMed=6945574;
RA Gelsler N., Weber K.;
RT "Comparison of the proteins of two immunologically distinct
intermediate-sized filaments by amino acid sequence analysis: desmin
and vimentin.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:4120-4123(1981).
-1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
-1- SUBCELLULAR LOCATION: CYTOPLASMIC.
-1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenses@isb-sib.ch).
CC -----
DR EMBL; AB011676; BAA25136.1; -
DR EMBL; AB011674; BAA25134.1; -
DR EMBL; AF136188; AAD46492.1; -

```

```

DR EMBL; AF363284; AAK51087.1; -
DR PIR; A02955; DMFG.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF; 1.
KM Intermediate filament; Coiled coil; Muscle protein.
FT INT_MET 0
FT DOMAIN 1 108 HEAD.
FT DOMAIN 109 412 ROD.
FT DOMAIN 413 470 TAIL.
FT DOMAIN 109 141 COIL_1A.
FT DOMAIN 142 151 LINKER_1.
FT DOMAIN 152 252 COIL_1B.
FT DOMAIN 253 268 LINKER_12.
FT DOMAIN 269 287 COIL_2A.
FT DOMAIN 288 295 LINKER_2.
FT DOMAIN 296 412 COIL_2B.
FT SITE 354 354 STUTTER.
FT SITE 44 47 POLY-SER.
SQ SEQUENCE 470 AA; 53497 MW; A2ABF7A8DB65DD12 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 1; Length 470;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKVHHE 6
DB 240 KKVHHE 245

RESULT 8
NFL_RAT STANDARD; PRT; 541 AA.
ID NFL_RAT 063367;
AC P19527; 063367;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NFL-L).
GN NFL OR NFL OR NF68.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90184052; PubMed=2516804;
RA Chin S.S., Liem R.K.H.;
RT "Expression of rat neurofilament proteins NF-L and NF-M in
transfected non-neuronal cells.";
RL Eur. J. Cell Biol. 50:475-490(1989).
RN [2]
RP SEQUENCE OF 197-483 FROM N.A.
RX MEDLINE=85252830; PubMed=3925999;
RA Jullen J.-P., Ramchandran K., Grosveld F.;
RT "Cloning of a cDNA encoding the smallest neurofilament protein from
the rat.";
RL Blochim. Biophys. Acta 825:398-404(1985).
RN [3]
RP SEQUENCE OF 1-10 FROM N.A.
RX MEDLINE=95264348; PubMed=745611;
RA Reegen M., Neuman T., Palgi J., Palm K., Paalme V., Saarma M.;
RT "Characterization of the rat light neurofilament (NF-L) gene promoter
and identification of NF and cAMP responsive regions.";
RL J. Neurosci. Res. 40:177-188(1995).
RN [4]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=93346421; PubMed=8344946;
RA Dong D.L.-Y., Xu Z.-S., Chevrier M.R., Cotter R.J., Cleveland D.W.,
RA Hart G.W.;
RT "Glycosylation of mammalian neurofilaments. Localization of multiple
O-linked N-acetylglucosamine moieties on neurofilament polypeptides
L and M.";

```

RL J. Biol. Chem. 268:16679-16687(1993).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- PTM: O-GLYCOSYLATED; CONTAINS THREE N-ACETYLGLUCOSAMINE SIDE
 CC CHAINS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF031880; AAB87069.1; -
 CC EMBL: M25638; AAA41694.1; -
 CC EMBL: X53981; CAA37931.1; -
 CC PIR: A21762; A21762.
 CC GlycoSuiteDB: P19527; -
 CC InterPro: IPR001664; IF.
 CC Pfam: PF00038; filament; 1.
 CC PROSITE: PS00226; IF; 1.
 CC DR Intermediate filament; Coiled coil; Neurone; Glycoprotein.
 CC KW INIT_MER 0 92
 CC FT DOMAIN 1 92
 CC FT DOMAIN 93 396
 CC FT DOMAIN 397 541
 CC FT DOMAIN 93 124
 CC FT DOMAIN 125 137
 CC FT DOMAIN 138 223
 CC FT DOMAIN 234 252
 CC FT DOMAIN 253 271
 CC FT DOMAIN 272 280
 CC FT DOMAIN 281 396
 CC FT DOMAIN 397 443
 CC FT DOMAIN 444 541
 CC FT CARBOHYD 20 20
 CC FT CARBOHYD 26 26
 CC FT SITE 381 391
 CC FT CONFLICT 197 202
 CC FT CONFLICT 399 399
 CC FT CONFLICT 476 476
 CC FT CONFLICT 480 483
 CC SQ SEQUENCE 541 AA; 61204 MW; 0D17839AF226918A CRC64;
 Query Match 100.0%; Score 32; DB 1; Length 541;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L).
 GN NEFL OR NEFL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RA MEDLINE=87064433; PubMed=3785173;
 RX Lewis S.A., Cowan N.J.;
 RT "Anomalous placement of introns in a member of the intermediate
 RT filament multigene family: an evolutionary conundrum";
 RL Mol. Cell. Biol. 6:1529-1534(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RA MEDLINE=87158637; PubMed=3103856;
 RX Julien J.-P., Meyer D., Flavell D., Hurst J., Grosved F.;
 RT "Cloning and developmental expression of the murine neurofilament
 RT gene family";
 RL Brain Res. 387:243-250(1986).
 RN [3]
 RP SEQUENCE OF 241-542 FROM N.A.
 RC TISSUE-Brain;
 RA MEDLINE=85131334; PubMed=3919033;
 RX Lewis S.A., Cowan N.J.;
 RT "Genetics, evolution, and expression of the 68,000-mol-wt
 RT neurofilament protein: isolation of a cloned cDNA probe";
 RL J. Cell Biol. 100:843-850(1985).
 RN [4]
 RP SEQUENCE OF 1-27 FROM N.A.
 RA MEDLINE=91060592; PubMed=2246261;
 RA Nakamura K., Ikenaka K., Kada K., Tamura T.A., Furuchi T.,
 RT Mikoshiba K.;
 RT "Structure of the 68-kDa neurofilament gene and regulation of its
 RT expression";
 RL J. Biol. Chem. 265:19786-19791(1990).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X02165; CAB51616.1; -
 CC EMBL: M20480; AAA39814.1; -
 CC EMBL: M13016; AAA39810.1; -
 CC EMBL: M54423; AAA39812.1; -
 CC PIR: A25227; OFNSL.
 CC MGI: 97313; NEFL.
 CC InterPro: IPR001664; IF.
 CC Pfam: PF00038; filament; 1.
 CC PROSITE: PS00226; IF; 1.
 CC KW Intermediate filament; Coiled coil; Neurone; Glycoprotein.
 CC FT INIT_MER 0 92
 CC FT DOMAIN 1 92
 CC FT HEAD.

FT DOMAIN 93 396 ROD.
 FT DOMAIN 542 124 TAIL.
 FT DOMAIN 93 124 COIL 1A.
 FT DOMAIN 125 137 LINKER 1.
 FT DOMAIN 138 233 COIL 1B.
 FT DOMAIN 234 252 LINKER 12.
 FT DOMAIN 253 271 COIL 2A.
 FT DOMAIN 272 280 LINKER 2.
 FT DOMAIN 281 396 COIL 2B.
 FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
 FT DOMAIN 444 542 TAIL, SUBDOMAIN B (ACIDIC).
 FT CARBOHYD 20 26 O-LINKED (GLCNAC) (BY SIMILARITY).
 FT CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
 FT SITE 391 391 EPIPTOPE (RECOGNIZED BY IF-SPECIFIC MONOCLONAL ANTIBODY).
 FT CONFLICT 5 5 Y -> S (IN REF. 1).
 FT CONFLICT 8 8 Y -> I (IN REF. 1).
 FT CONFLICT 64 64 M -> K (IN REF. 1).
 FT CONFLICT 72 72 V -> L (IN REF. 2).
 FT CONFLICT 98 98 D -> H (IN REF. 1).
 FT CONFLICT 194 194 R -> A (IN REF. 1).
 FT CONFLICT 202 202 MISSING (IN REF. 2).
 FT CONFLICT 239 239 Y -> I (IN REF. 1).
 SQ SEQUENCE 542 AA; 61448 MW; 8EE9B8C6F0831DBC CRC64;
 Query Match 100.0%; Score 32; DB 1; Length 542;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KKVHEE 6
 DB 223 KKVHEE 228
 RESULT 10
 ID NFL_HUMAN STANDARD; PRT; 543 AA.
 AC P07196; Q16154; Created
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L).
 GN NEFL OR NEFL OR NF68.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87214213; PubMed=3034332;
 RA Julien J.-P., Grosfeld F., Yazdankhah K., Flavell D., Melzer D.,
 RA Mushynski W.;
 RT "The structure of a human neurofilament gene (NF-L): a unique exon-
 RT intron organization in the intermediate filament gene family.";
 RL Biochim. Biophys. Acta 909:10-20(1987).
 RN [2]
 RP SEQUENCE OF 1-10 FROM N.A.
 RX MEDLINE=94235564; PubMed=8180132;
 RA Pospelova V.A., Pospelova T.V., Julien J.-P.;
 RA "Ap-1 and Krox-24 transcription factors activate the neurofilament
 RT light gene promoter in P19 embryonal carcinoma cells.";
 RL Cell Growth Differ. 5:187-196(1994).
 RN [3]
 RP VARIANT CMT2E PRO-332.
 RX MEDLINE=20307176; PubMed=10841809;
 RA Merisyanova I.V., Perepelov A.V., Polyakov A.V., Stukov V.F.,
 RA Padali E.L., Oparin R.B., Petrin A.N., Evgenov O.V.;
 RA "A new variant of Charcot-Marie-Tooth disease type 2 is probably the
 RT result of a mutation in the neurofilament-light gene.";
 RL Am. J. Hum. Genet. 67:37-46(2000).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.

CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- DISEASE: DEFECTS IN NEFL ARE A CAUSE OF CHARCOT-MARIE-TOOTH
 CC DISEASE TYPE 2E (CMT2E).
 CC -1- MISCELLANEOUS: NEFL IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NEONEUTRILIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -1- DATABASE: NAME-Inherited peripheral neuropathies mutation db;
 CC WWW="http://molgen-www.uia.ac.be/CMTmutations/".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.sdb.ch).
 CC -----
 CC DR EMBL; X05608; CAA29097.1; -;
 CC DR EMBL; S70309; AAD14057.1; -;
 CC DR PIR; S07144; S07144.
 CC DR Genew; HGNC:7739; NEFL.
 CC DR MIM; 162280;
 CC DR InterPro; IPR001664; IF.
 CC DR Pfam; PF00038; filament; 1.
 CC DR PROSITE; PS00226; IF; 1.
 CC DR Intermediate filament; Coiled coil; Neurone; Glycoprotein;
 CC Disease mutation; Charcot-Marie-Tooth disease.
 CC FT INT_MET 0 0
 CC FT DOMAIN 1 91 HEAD.
 CC FT DOMAIN 92 396 ROD.
 CC FT DOMAIN 397 543 TAIL.
 CC FT DOMAIN 92 123 COIL 1A.
 CC FT DOMAIN 124 136 LINKER 1.
 CC FT DOMAIN 137 234 COIL 1B.
 CC FT DOMAIN 235 252 LINKER 12.
 CC FT DOMAIN 253 271 COIL 2A.
 CC FT DOMAIN 272 280 LINKER 2.
 CC FT DOMAIN 281 396 COIL 2B.
 CC FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
 CC FT DOMAIN 444 543 TAIL, SUBDOMAIN B (ACIDIC).
 CC FT CARBOHYD 20 26 O-LINKED (GLCNAC) (BY SIMILARITY).
 CC FT CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
 CC FT SITE 381 391 EPIPTOPE (RECOGNIZED BY IF-SPECIFIC MONOCLONAL ANTIBODY).
 CC FT VARIANT 332 332 Q -> P (IN CMT2E).
 CC FT SEQUENCE 543 AA; 61645 MW; 7A0FIADDSB22F6 CRC64;
 SQ
 Query Match 100.0%; Score 32; DB 1; Length 543;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KKVHEE 6
 DB 224 KKVHEE 229
 RESULT 11
 ID NFL_XENLA STANDARD; PRT; 544 AA.
 AC P35616;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (Neurofilament light polypeptide) (NF-
 DE L).
 CC Xenopus laevis (African clawed frog).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 CC Xenopodidae; Xenopus.
 CC NCBI_TaxID=8355;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92356194; PubMed=1494944;
 RA Charnas L.R., Szaro B.G., Gainer H.;
 RT "Identification and developmental expression of a novel low molecular
 RT weight neuronal intermediate filament protein expressed in Xenopus
 RT laevis";
 RL J. Neurosci. 12:3010-3024(1992).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M86654; AAAB3018.1; -
 DR PIR: B44841; B44841.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; Filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neurene.
 FT DOMAIN 1 87 HEAD.
 FT FT DOMAIN 88 380 ROD.
 FT FT DOMAIN 381 544 TAIL.
 FT FT DOMAIN 88 119 COIL 1A.
 FT FT DOMAIN 120 132 LINKER 1.
 FT FT DOMAIN 133 228 COIL 1B.
 FT FT DOMAIN 229 246 LINKER 12.
 FT FT DOMAIN 247 265 COIL 2A.
 FT FT DOMAIN 266 274 LINKER 2.
 FT FT DOMAIN 275 390 COIL 2B.
 FT FT DOMAIN 391 435 TAIL, SUBDOMAIN A.
 FT FT DOMAIN 436 544 TAIL, SUBDOMAIN B (ACIDIC).
 FT FT DOMAIN 441 538 GLU-RICH.
 FT FT DOMAIN 464 469 POLY-GLU.
 SQ SEQUENCE 544 AA; 61861 MW; 76D911B896E97201 CRC64;
 Query Match 100.0%; Score 32; DB 1; Length 544;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KRVHEE 6
 DB 218 KRVHEE 223
 RESULT 12
 ID NFL_PIG STANDARD; PRT; 548 AA.
 AC P02547;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L).
 GN NEFL.
 OS Sus scrofa (Pig).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

CC NCBI_TaxID=9823;
 RN (1)
 RP SEQUENCE.
 RX MEDLINE=85154583; PubMed=3920075;
 RA Geisler N., Plessmann U., Weber K.;
 RT "The complete amino acid sequence of the major mammalian
 RT neurofilament protein (NF-L).";
 RL FEBS Lett. 182:475-478(1985).
 RN (2)
 RP SEQUENCE OF 1-82 AND 278-548.
 RA Geisler N., Kaufmann E., Fischer S., Plessmann U., Weber K.;
 RT "Neurofilament architecture combines structural principles of
 RT intermediate filaments with carboxy-terminal extensions increasing
 RT in size between triplet proteins";
 RL EMBO J. 2:1295-1302(1983).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 DR PIR: A09563; OEPGL.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; Filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neurene; Glycoprotein.
 FT DOMAIN 1 91 HEAD.
 FT FT DOMAIN 92 395 ROD.
 FT FT DOMAIN 396 548 TAIL.
 FT FT DOMAIN 92 123 COIL 1A.
 FT FT DOMAIN 124 136 LINKER 1.
 FT FT DOMAIN 137 232 COIL 1B.
 FT FT DOMAIN 233 251 LINKER 12.
 FT FT DOMAIN 252 270 COIL 2A.
 FT FT DOMAIN 271 279 LINKER 2.
 FT FT DOMAIN 280 395 COIL 2B.
 FT FT DOMAIN 396 442 TAIL, SUBDOMAIN A.
 FT FT DOMAIN 443 548 TAIL, SUBDOMAIN B (ACIDIC).
 FT FT CARBOHYD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
 FT FT CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
 FT FT SITE 380 390 EPITOPE (RECOGNIZED BY IF-SPECIFIC
 FT MONOCLONAL ANTIBODY).
 SQ UNSURE 322 322 OR K.
 SQ SEQUENCE 548 AA; 61940 MW; 83044813637AC739 CRC64;
 Query Match 100.0%; Score 32; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KRVHEE 6
 DB 222 KRVHEE 227
 RESULT 13
 ID NFL_BOVIN STANDARD; PRT; 554 AA.
 AC P02548; P79127;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L) (Micro glutamic acid-rich
 DE protein).
 GN NEFL.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovine; Bos.
 CC NCBI_Taxid=9913;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-Holstein; TISSUE-Brain;
 RA Hill W.D., Zhang L., Ballin B.J., Sprinkle T.J.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE OF 468-554.
 RX MEDLINE=85154567; PubMed=3884373;
 RA "Brain micro glutamic acid-rich protein is the C-terminal endpiece of
 RT the neurofilament 68-kDa protein as determined by the primary
 RT sequence."
 RL Febs Lett. 182:389-392(1985).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U83919; ABA41543.1; -
 DR PIR: A02964; OPRO
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KM Intermediate filament; Colled coil; Neutrone.
 FT INIT_MET 0 0
 FT DOMAIN 1 92 HEAD (BY SIMILARITY).
 FT DOMAIN 93 396 ROD (BY SIMILARITY).
 FT DOMAIN 397 554 TAIL (BY SIMILARITY).
 FT DOMAIN 125 137 COIL 1A.
 FT DOMAIN 138 233 COIL 1B.
 FT DOMAIN 234 252 COIL 12.
 FT DOMAIN 253 271 COIL 2A.
 FT DOMAIN 272 280 COIL 2B.
 FT DOMAIN 281 396 LINKER 1.
 FT DOMAIN 397 443 LINKER 12.
 FT DOMAIN 444 554 TAIL, SUBDOMAIN A.
 FT DOMAIN 494 500 TAIL, SUBDOMAIN B (ACIDIC).
 FT CONFLICT 509 509 MISSING (IN REF. 2).
 FT CONFLICT 509 509 A > AEA (IN REF. 2).
 SQ SEQUENCE 554 AA; 62514 MW; D772B81CA2C31CIA CRC64;
 QY 1 KVVHEE 6
 DB 223 KVVHEE 228
 Query Match 100.0%; Score 32; DB 1; Length 554;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DT 01-FEB-1994 (rel. 28, Created)
 DT 01-JUN-1994 (rel. 29, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (Neurofilament light polypeptide) (NF-
 DE L).
 GN NEFL.
 OS Coturnix coturnix japonica (Japanese quail).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Coturnix.
 CC NCBI_Taxid=93934;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=93224534; PubMed=8468353;
 RA Ohara O., Gahara Y., Miyake T., Teraoka H., Kitamura T.;
 RT "Neurofilament deficiency in quail caused by nonsense mutation in
 RT neurofilament-L gene."
 RL J. Cell Biol. 121:387-395(1993).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- DISEASE: NF-L DEFICIENCY CAUSES THE DISORDER QUIVER.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D13223; BAA02504.1; -
 DR EMBL: D13222; BAA02503.1; ALT_TERM.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KM Intermediate filament; Colled coil; Neutrone.
 FT INIT_MET 0 0
 FT DOMAIN 1 93 HEAD (BY SIMILARITY).
 FT DOMAIN 94 397 ROD (BY SIMILARITY).
 FT DOMAIN 398 555 TAIL (BY SIMILARITY).
 FT DOMAIN 125 137 COIL 1A.
 FT DOMAIN 138 233 COIL 1B.
 FT DOMAIN 234 252 COIL 12.
 FT DOMAIN 253 271 COIL 2A.
 FT DOMAIN 272 281 COIL 2B.
 FT DOMAIN 282 397 LINKER 1.
 FT DOMAIN 398 444 LINKER 12.
 FT DOMAIN 445 555 TAIL, SUBDOMAIN A.
 FT CONFLICT 555 555 TAIL, SUBDOMAIN B (ACIDIC).
 SQ SEQUENCE 555 AA; 62282 MW; 9B957ABDB8BA7712 CRC64;
 QY 1 KVVHEE 6
 DB 224 KVVHEE 229
 Query Match 100.0%; Score 32; DB 1; Length 555;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14
 ID NFL_COTJA
 AC 002916;
 STANDARD; PRT; 555 AA.

RESULT 15
 ID E2BL_METJA
 E2BL_METJA


```

ID  E2B1 METJA      STANDARD;      PRT;      329 AA.
AC  057896;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Putative translation initiation factor eIF-2B subunit 1 (eIF-2B
GN  GDP-GTP exchange factor).
OS  M00454.
OS  Methanococcus jannaschii.
OC  Archaea: Euryarchaeota: Methanococci: Methanococcales;
OC  Methanocaldococcaceae; Methanocaldococcus.
OX  NCBI_TaxID=2190;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX  MEDLINE=96337999; PubMed=8688087;
RA  Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA  Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA  Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA  Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA  Scott J.L., Geoghegan N.S.M., Feldman J.F., Fuhrmann J.L., Nguyen D.,
RA  Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA  Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA  Klepek H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT  "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT  jannaschii."
RL  Science 273:1058-1073(1996).
CC  -1- FUNCTION: CATALYZES THE EXCHANGE OF INITIATION FACTOR 2-BOUND GDP
CC  FOR GTP (BY SIMILARITY).
CC  -1- SUBUNIT: COMPLEX OF TWO DIFFERENT SUBUNITS (POTENTIAL).
CC  -1- SIMILARITY: BELONGS TO THE EIF-2B ALPHA/BETA/DELTA SUBUNITS
CC  FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL: U67496; AAB98443.1; .
DR  TIGR: M00454; .
DR  InterPro: IPR005251; AIF-2B1_fam.
DR  InterPro: IPR000649; IF-2B.
DR  Pfam: PF01008; IF-2B_1.
DR  TIGRFAMS: TIGR00512; aif-2B1_fam; 1.
DR  TIGRFAMS: TIGR00524; eif-2B_rel; 1.
KW  Hypothetical protein; Initiation factor; Protein biosynthesis;
KW  Complete proteome.
SQ  SEQUENCE 329 AA; 36802 MW; 3936298C574730C9 CRC64;

```

```

Query Match          96.9%; Score 31; DB 1; Length 329;
Best Local Similarity 83.3%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY  1 KKVHEE 6
    11:1111
DB  112 KKVHEE 117

```

Search completed: November 13, 2002, 13:16:53
 Job time : 5.97872 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 19.2766 Seconds

(without alignments)
64.134 Million cell updates/sec

Title: US-09-856-086-5

Perfect score: 32

Sequence: 1 KKVHEE 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP-archaea:*
2: SP-bacteria:*
3: SP-fungi:*
4: SP-human:*
5: SP-invertebrate:*
6: SP-mammal:*
7: SP-mhc:*
8: SP-organellar:*
9: SP-phage:*
10: SP-plant:*
11: SP-rodent:*
12: SP-virus:*
13: SP-vertebrate:*
14: SP-unclassified:*
15: SP-virus:*
16: SP-bacteriap:*
17: SP-archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	32	100.0	60 13 Q91377	Q91377 gallus gall
2	32	100.0	68 13 Q91355	Q91355 gallus gall
3	32	100.0	73 4 Q9GZ86	Q9GZ86 homo sapien
4	32	100.0	144 10 Q9C651	Q9C651 arabidopsis
5	32	100.0	386 4 Q8RCR7	Q8RCR7 homo sapien
6	32	100.0	448 13 Q73665	Q73665 gallus gall
7	32	100.0	455 13 Q73587	Q73587 scyllorhinu
8	32	100.0	470 4 Q9UHN5	Q9UHN5 homo sapien
9	32	100.0	470 4 Q9H319	Q9H319 homo sapien
10	32	100.0	470 4 Q8RD99	Q8RD99 homo sapien
11	32	100.0	470 4 Q8RCX2	Q8RCX2 homo sapien
12	32	100.0	470 4 Q8RCX4	Q8RCX4 homo sapien
13	32	100.0	501 11 Q9CAX8	Q9CAX8 mus musculu
14	32	100.0	1937 11 Q91BD4	Q91BD4 gallus gall
15	31	96.9	241 10 Q9LVE7	Q9LVE7 arabidopsis
16	31	96.9	410 5 Q18665	Q18665 caenorhabdit

17	31	96.9	459 13 Q8WPF1	Q8WPF1 oncorhynch
18	31	96.9	807 10 Q48724	Q48724 arabidopsis
19	31	96.9	958 4 Q8WMI2	Q8WMI2 homo sapien
20	31	96.9	3134 5 Q25594	Q25594 plasmodium
21	31	96.9	3144 5 Q9GTR4	Q9GTR4 plasmodium
22	29	90.6	75 17 Q96ZV9	Q96ZV9 sulfobolus
23	29	90.6	146 16 Q8XKN5	Q8XKN5 clostridium
24	29	90.6	197 17 Q8T882	Q8T882 methanopyru
25	29	90.6	204 2 Q91723	Q91723 vibrio prot
26	29	90.6	216 16 Q9X0Y1	Q9X0Y1 thermotoga
27	29	90.6	262 16 Q9TEW4	Q9TEW4 clostridium
28	29	90.6	265 12 Q9QJ13	Q9QJ13 human herpe
29	29	90.6	270 12 Q57159	Q57159 human herpe
30	29	90.6	275 4 Q96H12	Q96H12 homo sapien
31	29	90.6	275 11 Q9CY55	Q9CY55 mus musculu
32	29	90.6	275 11 Q8CR78	Q8CR78 mus musculu
33	29	90.6	283 17 Q8ZWS1	Q8ZWS1 pyrobaculum
34	29	90.6	333 10 Q9ZQ88	Q9ZQ88 arabidopsis
35	29	90.6	333 16 Q67652	Q67652 aquifex aeo
36	29	90.6	349 16 Q8RDG7	Q8RDG7 thermomane
37	29	90.6	364 16 Q9HXY7	Q9HXY7 pseudomonas
38	29	90.6	387 10 Q9FX81	Q9FX81 arabidopsis
39	29	90.6	410 4 Q96ML2	Q96ML2 homo sapien
40	29	90.6	450 13 Q9DD83	Q9DD83 scyllorhinu
41	29	90.6	451 17 Q96ZG0	Q96ZG0 sulfobolus
42	29	90.6	459 13 Q8UVG8	Q8UVG8 dabola rurs
43	29	90.6	471 13 B87360	B87360 brachydanio
44	29	90.6	472 13 Q07962	Q07962 carassius a
45	29	90.6	475 4 Q9H8F5	Q9H8F5 homo sapien

ALIGNMENTS

RESULT 1
ID Q91377 PRELIMINARY: PRT: 60 AA.
AC Q91377;
DI 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-NOV-2001 (TREMBLrel. 17, Last annotation update)
DE Myosin heavy chain (fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1].
RP SEQUENCE FROM N.A.
RX MEDLINE=94227840; PubMed=8173588;
RA Camarotti-Mercado B., Dizon E., Jakovcic S., Zak R.;
RT "Differential expression of ventricular-like myosin heavy chain mRNA
in developing and regenerating avian skeletal muscles.";
RL Cell. Mol. Biol. Res. 39:425-437(1993).
DR EMBL: S70443; AAB30704.1; -;
DR InterPro: IPR002928; Myosin_tail.
DR Pfam: PF01576; Myosin_tail; 1.
FT NON-TER
SQ SEQUENCE 60 AA; 7060 MW; 3165EE04EAE24AE1 CRC64;
FT NON-TER
Query Match 100.0%; Score 32; DB 13; Length 60;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KKVHEE 6
DB 54 KKVHEE 59
RESULT 2
ID Q91355 PRELIMINARY: PRT: 68 AA.
AC Q91355;

```

DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Ventricular myosin heavy chain 1 (Fragment).
OS Gallus gallus (chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92038448; PubMed=1936571;
RA Bisaha J.G., Bader D.;
RT "Identification and characterization of a ventricular-specific avian
RT myosin heavy chain, VMHC1: expression in differentiating cardiac and
RT skeletal muscle."
RL Dev. Biol. 148:355-364(1991).
DR EMBL; S64689; AAB20271.1; -.
DR InterPro; IPR002928; Myosin_tail.
DR Pfam; PF01576; Myosin_tail; 1.
FT NON_TER
SO SEQUENCE 68 AA; 8013 MW; 8565E1CBACB4E47 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 13; Length 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 62 KKVHEE 67

RESULT 3
O9GZRG PRELIMINARY; PRT; 73 AA.
AC O9GZRG;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Mutant desmin CSM-7 (Mutant desmin CSM-6) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Park K.-Y., Dalakas M.C., Goebel H.H., Ferrans V.J., Semino-Mora C.,
RA Litvak S., Takeda K., Goldfarb L.G.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF189281; AAC3460.1; -.
DR EMBL; AF189280; AAC3459.1; -.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; Filament; 1.
FT NON_TER
FT NON_TER
SO SEQUENCE 73 AA; 8537 MW; B062E7B4455FE955 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 4; Length 73;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 47 KKVHEE 52

RESULT 4
O9G651 PRELIMINARY; PRT; 144 AA.
AC O9G651;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

```

```

DE Hypothetical 16.4 kDa protein.
GN F9K23.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA.
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.-J., Koo H.-L., Kremetska I., Kurtz D.B., Kvan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marzilli A.,
RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utechtack T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
DR EMBL; AC082643; AAG50642.1; -.
RW Hypothetical protein.
SO SEQUENCE 144 AA; 16438 MW; 5D2A4A26576131DF CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 10; Length 144;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 89 KKVHEE 94

RESULT 5
O8TCR7 PRELIMINARY; PRT; 386 AA.
AC O8TCR7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 43.8 kDa protein (Fragment).
GN DKF2P761K0922.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=AMYGDA;
RA Wandut R., Heubner D., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL713644; CAD28456.1; -.
RW Hypothetical protein.
FT NON_TER
FT NON_TER
SO SEQUENCE 386 AA; 43779 MW; F293388B200C7B65 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 4; Length 386;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 67 KKVHEE 72

```

```

RESULT 6
ID 073665 PRELIMINARY; PRT; 448 AA.
AC 073665;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Desmin (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-WHITE LECHORN; TISSUE=MUSCLE;
RA Chikuni K.;
RT "Partial sequence of chicken desmin."
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
DR EMBL; AF011672; BA25132.1; -.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF. 1.
KW Coiled coil; Intermediate filament.
FT NON-TER 1
FT NON-TER 448
SQ SEQUENCE 448 AA; 51664 MW; 198BEC377DB94B40 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 13; Length 448;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 223 KKVHEE 228

RESULT 7
ID 073587 PRELIMINARY; PRT; 455 AA.
AC 073587;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Desmin.
OS Scyliorhinus stellaris.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeidae; Carcharhiniformes;
OC Scyliorhinidae; Scyliorhinus.
OX NCBI_TaxID=68454;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Schultess J.; Loebbecke A.; Schaffeld M.; Lieb B.; Markl J.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Schaffeld M.; Schultess J.; Loebbecke A.; Lieb B.; Herrmann H.;
RA Markl J.;
RT "Primary structure, expression patterns and properties of vimentin and
desmin in the shark Scyliorhinus stellaris."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
DR EMBL; Y15064; CA475347.1; -.
DR EMBL; AJ304374; CAC83054.1; -.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF. 1.
KW Coiled coil; Intermediate filament.

```

```

SQ SEQUENCE 455 AA; 52230 MW; B330A2FCF895BCE9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 13; Length 455;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 224 KKVHEE 229

RESULT 8
ID 09UNH5 PRELIMINARY; PRT; 470 AA.
AC 09UNH5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Mutant desmin.
GN MUTDESM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99359868; PubMed=10430757;
RA Li D.X., Tapscott T., Gonzalez O., Burch P.E., Quinones M.A.,
RA Zoghbi W.A., Hill R., Bachisk L.L., Mann D.L., Roberts R.;
RT "Desmin Mutation Responsible for Dilated Cardiomyopathy."
RL Circulation 100:461-464(1999).
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
DR EMBL; AF137053; AAF15400.1; -.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF. 1.
KW Coiled coil; Intermediate filament.
FT VARIANT 451
FT VARIANT 451 M->I.
SQ SEQUENCE 470 AA; 53553 MW; 1B5D9EAE51FBB319 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 4; Length 470;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 240 KKVHEE 245

RESULT 9
ID 09H319 PRELIMINARY; PRT; 470 AA.
AC 09H319;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Mutant desmin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Park K.-Y., Dalakas M.C., Semino-Mora C., Lee H.-S., Litvak S.,
RA Goldfarb L.G.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF167579; AAG41217.1; -.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
FT VARIANT 406
FT VARIANT 406 W->R.
SQ SEQUENCE 470 AA; 53565 MW; 1C2AEAE4B4BB319 CRC64;

Query Match
Best Local Similarity 100.0%; Score 32; DB 4; Length 470;

```

Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 240 KKVHEE 245

RESULT 10

O8TD99 PRELIMINARY; PRT; 470 AA.

AC O8TD99;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Mutant desmIn.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Degvadof A.;
RT "Novel desmIn mutation causing complete disruption of the intermediate filament network."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF487828; AAL9078.1; -
SQ SEQUENCE 470 AA; 53561 MW; BBS1286FE75605DF CRC64;

Query Match Best Local Similarity 100.0%; Score 32; DB 4; Length 470;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 240 KKVHEE 245

RESULT 11

O8TCX2 PRELIMINARY; PRT; 470 AA.

AC O8TCX2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Mutant desmIn.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Degvadof A.;
RT "Proline substitution in mutant desmIn breaks the alpha-helical conformation and results in severe phenotype."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY083345; AAL9215.1; -
SQ SEQUENCE 470 AA; 53519 MW; OEB5F052D6FE23C CRC64;

Query Match Best Local Similarity 100.0%; Score 32; DB 4; Length 470;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 240 KKVHEE 245

RESULT 12

O8TAC4 PRELIMINARY; PRT; 470 AA.

AC O8TAC4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Mutant desmIn.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21523003; PubMed=11668632;
RA Goudreau B., Degvadof A., Rodrigues-Lima F., Nedeliec P.,
RA Casters-Simon M., Perret E., Langlois S., Goldfarb L., Vicart P.;
RT "Structural and functional analysis of a new desmIn variant causing Hum. Mutat. 18:388-396(2001)."
RN [2]
RP SEQUENCE FROM N.A.
RA Goudreau B., Degvadof A., Rodrigues-Lima F., Nedeliec P.,
RA Casters-Simon M., Perret E., Langlois S., Goldfarb L., Vicart P.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF486807; AAL93205.1; -
SQ SEQUENCE 470 AA; 53504 MW; 1B5C44024DE11869 CRC64;

Query Match Best Local Similarity 100.0%; Score 32; DB 4; Length 470;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
DB 240 KKVHEE 245

RESULT 13

O9CX98 PRELIMINARY; PRT; 501 AA.

AC O9CX98;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE 8430436A10R1K protein.
GN 8430436A10R1K.
CN 8430436A10R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRIONIC LUNG;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batlov S., Casavant T.,
RA Fletschmann W., Gaasterland T., Gissi C., King B., Kocha H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pescio G., Quackenbush J.,
RA Sakurai L.M., Steubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamalya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombert P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -1 SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AK018458; BAB31223.1; -
DR HSSP; P00179; 1D76.
MGD; MGI:1918769; 8430436A10R1K.

DR InterPro: IPR001128; Cytochrome_P450.
 DR Pfam: PF00067; P450.1.
 DR PRINTS: PR00385; P450.
 KW Heme; Monooxygenase; Oxidoreductase.
 SQ SEQUENCE 501 AA; 57320 MW; 4A409AA7E463D369 CRC64;

Query Match 100.0%; Score 32; DB 11; Length 501;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 |||||
 DB 362 KKVHEE 367

RESULT 14

O9IBD4 PRELIMINARY; PRT; 1937 AA.

AC O9IBD4; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Myosin heavy chain.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WHITE LEGHORN; TISSUE=HEART;
 RA MEDLINE=2014854; PubMed=10684978;
 RX Machida S., Noda S., Furutani Y., Takao A., Momma K., Matsuka R.;
 RT "Complete sequence and characterization of chick ventricular myosin
 heavy chain in the developing atria."
 RL Blochum. Biophys. Acta 1490:333-341(2000).
 DR EMBL; AB032197; BAA92710.1; -
 DR HSBP; P08799; IAMD.
 DR InterPro; IPR000048; IQ_region.
 DR InterPro; IPR001609; myosin_head.
 DR InterPro; IPR004009; myosin_N.
 DR InterPro; IPR002928; myosin_tail.
 DR InterPro; IPR000533; Tropomyosin.
 DR Pfam; PF00612; IQ.2.
 DR Pfam; PF00063; myosin_head.1.
 DR Pfam; PF02736; myosin_N.1.
 DR Pfam; PF01576; myosin_tail.1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR PRINTS; PR00194; TROPOMYOSIN.
 DR PRODOM; PD000355; myosin_head.1.
 DR SMART; SM00015; IQ.1.
 DR SMART; SM00242; MYSC.1.
 SQ SEQUENCE 1937 AA; 223473 MW; 95A94F69CD836781 CRC64;

Query Match 100.0%; Score 32; DB 13; Length 1937;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 |||||
 DB 1937 KKVHEE 1936

RESULT 15

O9LVE7 PRELIMINARY; PRT; 241 AA.

AC O9LVE7; 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE EmbiCAB45990.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA.
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
 RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA.
 RA MEDLINE=20277480; PubMed=10819329;
 RX Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 features of the regions of 4,504,864 bp covered by sixty P1 and TAC
 clones."
 RL DNA Res. 7:131-135(2000).
 DR EMBL; AB019232; BAB02354.1; -
 SQ SEQUENCE 241 AA; 25848 MW; F90851E5592DCE9 CRC64;

Query Match 96.9%; Score 31; DB 10; Length 241;
 Best Local Similarity 83.3%; Pred. No. 63;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 |||||
 DB 115 KKVHEE 120

Search completed: November 13, 2002, 13:16:09
 Job time : 20.2766 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 29.7447 Seconds
(without alignments)
26.879 Million cell updates/sec

Title: US-09-856-086-6

Perfect score: 29

Sequence: 1 EIRDLR 6

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.GeneSeq_101002:*

1: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*

2: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*

3: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*

4: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*

5: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*

6: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*

7: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*

8: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*

9: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*

10: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*

11: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*

12: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*

13: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*

14: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*

15: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*

16: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*

17: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*

18: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*

19: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*

20: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*

21: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*

22: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*

23: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	29	100.0	6	21	AA05931	Test antigen #7 fo
2	29	100.0	8	21	AA05932	Test antigen #8 fo
3	29	100.0	441	19	AA120612	Human neurofilamen
4	29	100.0	543	22	AA033466	Human polypeptide,
5	28	96.6	108	22	AA063810	Propionibacterium
6	28	96.6	338	21	AA020771	Arabidopsis thalia
7	28	96.6	338	21	AA037284	Arabidopsis thalia
8	28	96.6	338	23	AB082267	Hericicidally activ
9	28	96.6	403	23	AB057243	Mouse ischemic co
10	28	96.6	1017	21	AA038597	Arabidopsis thalia

11	27	93.1	265	23	AB072331
12	27	93.1	299	22	AB049729
13	27	93.1	323	23	AB059541
14	27	93.1	554	23	AB027632
15	27	93.1	555	22	AB031943
16	27	93.1	556	22	AB061625
17	27	93.1	556	23	AB027633
18	27	93.1	563	23	AB061631
19	27	93.1	574	22	AB029271
20	27	93.1	579	23	AB047410
21	27	93.1	580	21	AA093274
22	27	93.1	585	22	AA033477
23	27	93.1	586	22	AA034987
24	27	93.1	790	21	AA026956
25	27	93.1	790	21	AA051603
26	27	93.1	796	23	AA050849
27	27	93.1	814	16	AA011940
28	26	89.7	14	22	AA068744
29	26	89.7	77	22	AA054409
30	26	89.7	87	22	AA018600
31	26	89.7	88	18	AA055461
32	26	89.7	89	18	AA055581
33	26	89.7	90	19	AA098624
34	26	89.7	90	19	AA071550
35	26	89.7	90	22	AA046346
36	26	89.7	90	22	AA046408
37	26	89.7	96	18	AA055192
38	26	89.7	148	23	AB000247
39	26	89.7	193	22	AB067199
40	26	89.7	220	19	AA086074
41	26	89.7	305	21	AA041899
42	26	89.7	325	22	AA098297
43	26	89.7	393	22	AA017125
44	26	89.7	418	22	AB024673
45	26	89.7	446	22	AB013153

ALIGNMENTS

RESULT 1

AA05931

AA05931 standard; peptide; 6 AA.

AA05931;

17-OCT-2000 (first entry)

Test antigen #7 for spongiform and demyelinating disease diagnosis.

Human: cow; myelin; myelin neurofilament; immunogen; antigen;

bovine spongiform encephalopathy; BSE; multiple sclerosis;

Creutzfeldt-Jacob disease; CJD; demyelinating disease; diagnostic test.

Bos taurus.

Homo sapiens.

WO200031545-A1.

02-JUN-2000.

25-NOV-1999; 99MO-GB03936.

26-NOV-1998; 98GB-0025948.

(UNLO) KING'S COLLEGE.

Ehringer A;

WPI; 2000-400194/34.

Diagnosing spongiform or demyelinating disease in vertebrates such as

bovine spongiform encephalopathy and Creutzfeldt-Jacob disease comprises

PT assaying a biological sample for myelin and/or myelin neurofilament
 PT antibodies -
 XX
 PS Claim 5; Page 2; 16pp; English.
 XX
 CC The present peptide may be used as a test antigen in a kit for diagnosing
 CC spongiform or demyelinating disease in vertebrates, including bovine
 CC spongiform encephalopathy (BSE), multiple sclerosis (MS) and
 CC Creutzfeld-Jacob disease (CJD). Peptides used in the kit are antigenic
 CC components of myelin or myelin neurofilaments. Biological samples
 CC are assayed for antibodies, especially IGA antibodies, which bind to
 CC myelin and/or myelin neurofilaments or an antigenic peptide, such as the
 CC present sequence. Any reading in excess of two standard deviations of
 CC the healthy controls would indicate a positive response.
 SO Sequence 6 AA:
 OY 1 EIRDLR 6
 111111
 Db 1 EIRDLR 6
 RESULT 2
 ID AAB05932
 AAB05932 standard; peptide: 8 AA.
 AC AAB05932;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Test antigen #8 for spongiform and demyelinating disease diagnosis.
 XX
 KM Human; cow; myelin; myelin neurofilament; immunogen; antigen;
 KM bovine spongiform encephalopathy; BSE; multiple sclerosis;
 KM Creutzfeld-Jacob disease; CJD; demyelinating disease; diagnostic test.
 XX
 OS Bos taurus.
 OS Homo sapiens.
 OS
 PN WO200031545-A1.
 PN
 PD 02-JUN-2000.
 XX
 PF 25-NOV-1999; 99WO-GB03936.
 XX
 PR 26-NOV-1998; 98GB-0025948.
 XX
 PA (UNLO) KING'S COLLEGE.
 PI
 PI Edlinger A;
 XX
 DR WPI; 2000-400194/34.
 XX
 PT Diagnosing spongiform or demyelinating disease in vertebrates such as
 PT bovine spongiform encephalopathy and Creutzfeld-Jacob disease comprises
 PT assaying a biological sample for myelin and/or myelin neurofilament
 PT antibodies -
 PT
 PS Claim 5; Page 2; 16pp; English.
 XX
 CC The present peptide may be used as a test antigen in a kit for diagnosing
 CC spongiform or demyelinating disease in vertebrates, including bovine
 CC spongiform encephalopathy (BSE), multiple sclerosis (MS) and
 CC Creutzfeld-Jacob disease (CJD). Peptides used in the kit are antigenic
 CC components of myelin or myelin neurofilaments. Biological samples
 CC are assayed for antibodies, especially IGA antibodies, which bind to
 CC myelin and/or myelin neurofilaments or an antigenic peptide, such as the
 CC present sequence. Any reading in excess of two standard deviations of
 CC the healthy controls would indicate a positive response.
 CC

XX	sequence	8 AA;
XX	Query Match	100.0%; Score 29; DB 21; Length 8;
XX	Best Local Similarity	100.0%; Pred. No. 7.Be+05;
XX	Matches 6; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
OY	1 EIRDLR 6	
Dd		
	3 EIRDLR 8	
 RESULT 3 AAAY20612 ID. AAY20612 standard; Protein; 441 AA. XX AC AAY20612; XX DT 22-JUL-1999 (first entry) DE Human neurofilament-L wild type protein fragment 2. XX XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer; KM frameshift mutation; age-related disease; neurodegenerative disorder; KM Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal; KM Huntington's disease; multiple sclerosis; alcoholic liver disease; KM diabetes mellitus type II; microtubule associated protein; Tau; Big Tau; KM ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M; KM neurofilament-F; presenilin I; presenilin II; cellular tumour antigen; KM g1a1 fibrillary acidic protein; GFAP; p53; semaphorin II; HUPF-I; KM bcl-2; B-cell leukemia/Lymphoma 2 proto-oncogene; HMGP-C; NSP-A; KM high mobility group protein-C; neuroendocrine specific protein A. XX OS Homo sapiens. XX PN WO9845322-A2. XX PD 15-OCT-1998. XX PF 02-APR-1998; 98MO-IB00705. XX PI 10-APR-1997; 97US-0043163. XX PA (UYRT-) RIJKSONIV UTRECHT. PA (ROYA-) ROYAL NETHERLANDS ACAD ARMS & SCI. PA (UYRO-) UNIV ROTTERDAM ERASMUS. PI Burdach JPH, Grosveld FG, Van Leeuwen FW; PI WPI: 1998-609901/51. XX N-PSTDB; AAX75758. DR PT Diagnosing disease by detecting frameshift mutations in RNA or PT corresponding protein mutations - used to diagnose cancer and PT neurological diseases, particularly Alzheimer's disease, and also PT for treatment and prevention with specific ribozymes or wild-type PT RNA XX XX Disclosure; Figure 7; 258pp; English. XX XX This invention describes a novel method for the diagnosis of a disease CC caused by, or associated with, an RNA molecule that has a frameshift CC mutation. The method is used to diagnose age-related diseases, especially CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II CC and many others listed) or susceptibility to these disorders. The method CC allows a definitive diagnosis of Alzheimer's disease in living patients, CC at an early stage. It is based on the observation that disease may be CC caused by mutations in RNA rather than DNA. The invention describes the CC use of neuronal system RNA molecules, specifically proteins including CC beta-amyloid precursor protein (beta-APP), the microtubule associated CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule		

CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

CC Sequence 441 AA;

Query Match 100.0%; Score 29; DB 19; Length 441;

Best Local Similarity 100.0%; Pred. No. 3.2e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0;

QY 1 EIRDLR 6

DB 179 EIRDLR 184

RESULT 4

AA093466 standard; Protein; 543 AA.

AC AA093466;

DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 3132.

KM Human; full length cDNA; cDNA synthesis; oligo-capping.

OS Homo sapiens.

PN EPI30094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99AP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000MP-0183765.

PA (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI: 2001-524255/58.

DR N-PSDB: AAK94387.

PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -

PS Claim 8; SEQ ID NO 3132; 1380bp + sequence listing; English.

CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been

CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by

CC length enriched cDNA libraries that were synthesized by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily

CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

SO Sequence 543 AA;

Query Match 100.0%; Score 29; DB 22; Length 543;

Best Local Similarity 100.0%; Pred. No. 4e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0;

QY 1 EIRDLR 6

DB 142 EIRDLR 147

RESULT 5

AA063810 standard; Protein; 108 AA.

AC AA063810;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #24706.

KM SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;
 KM uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;

KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KM dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US12865.

PR 21-APR-2000; 2000US-199047P.

PR 02-JUN-2000; 2000US-208841P.

PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.

PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'walsonneuve J, Zhang Y, Jen S, Carter D;

DR WPI: 2001-616774/71.

DR N-PSDB: AAS59635.

PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -

PS Example 1; SEQ ID NO 25005; 1069bp; English.

CC Sequences AA093105-AA068017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in

CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory

CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a

CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The

CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to

CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as

CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 108 AA;

Query Match 96.6%; Score 28; DB 22; Length 108;

Best Local Similarity 83.3%; Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1;

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0155659.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158329.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 96.6%; Score 28; DB 21; Length 338;
Best Local Similarity 83.3%; Pred. No. 4e+02; 0; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0;

OY 1 EIRDLR 6
DB 93 EVRDLR 98

RESULT 7
AAG37284

ID AAG37284 standard; Protein; 338 AA.
XX
AC AAG37284;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45620.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0138461.
PR 18-JUN-1999; 99US-0138462.
PR 18-JUN-1999; 99US-0138463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144633.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150584.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153700.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0158295.
PR 14-OCT-1999; 99US-0158329.
PR 14-OCT-1999; 99US-0158330.
PR 14-OCT-1999; 99US-0158331.
PR 14-OCT-1999; 99US-0158637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 29-OCT-1999; 99US-0162142.

Query Match 96.6%; Score 28; DB 21; Length 338;
Best Local Similarity 83.3%; Pred. No. 4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ETRDLR 6
Db 93 EYRDLR 98

RESULT 8
ABB92267
ID ABB92267 standard; Protein: 338 AA.

XX ABB92267;
AC ABB92267;
XX 31-MAY-2002 (first entry)
DT
XX Herbidically active polypeptide SEQ ID NO 1478.

```

XX Herbicidal; plant; agriculture; herbicide.
KW Arabidopsis thaliana.
OS
XX WO200210210-A2.
PN
XX 07-FEB-2002.
PD
XX 28-AUG-2001; 2001WO-EP09892.
PF
XX 28-AUG-2001; 2001WO-EP09892.
PR
XX (FARB ) BAYER AG.
PA
XX Tietjen K, Weidner M;
PI
XX WPI; 2002-269010/31.
DR
XX
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
XX Claim 5; SEQ ID NO 1478; 261pp + Sequence Listing; English.
PS
XX
XX The invention relates to identifying target proteins
CC (AB990790-AB994016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
CC
XX Sequence 338 AA;
SO
XX
XX Query Match 96.6%; Score 28; DB 23; Length 338;
XX Best Local Similarity 83.3%; Pred. No. 4e+02;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 EIRDLR 6
XX 1:||||
DB 93 EVRDLR 98
XX
XX RESULT 9
XX ABB57243
XX ID ABB57243 standard; Protein: 403 AA.
XX
XX ABB57243;
AC
XX
XX 07-MAR-2002 (first entry)
DT
XX
XX Mouse ischaemic condition related protein sequence SEQ ID NO:650.
DE
XX
XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW vasoospastic ischaemia; ischaemic condition; ischaemic disease.
XX
XX Mus musculus.
OS
XX WO200186188-A2.
PN
XX
XX 22-NOV-2001.
PD
XX
XX 18-MAY-2001; 2001WO-JP04192.
PF
XX 18-MAY-2000; 2000JP-0145977.
PR
XX
XX (DVNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
PA
XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
PI

```

```

XX WPI; 2002-034733/04.
DR N-PSDB; AB199627.
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
XX Claim 2; Page 1635-1637; 2690pp; English.
PS
XX
XX The present invention describes a method for examining ischemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasoospastic ischaemia) by measuring
CC expression levels of particular genes (AB199202 to AB199912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischemic condition-improving
CC drugs or therapeutics for ischemic diseases. AB199913 and AB199914
CC represent PCR primers for a mouse ischemic condition related sequence,
CC which are used in the exemplification of the present invention.
CC
XX Sequence 403 AA;
SO
XX
XX Query Match 96.6%; Score 28; DB 23; Length 403;
XX Best Local Similarity 83.3%; Pred. No. 4.8e+02;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 EIRDLR 6
XX 1:||||
DB 177 EVRDLR 182
XX
XX RESULT 10
XX AAG38597
XX ID AAG38597 standard; Protein: 1017 AA.
XX
XX AAG38597;
AC
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 47640.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.

```

PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132040.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132861.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139765.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142350.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148317.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153750.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160880.
 PR 22-OCT-1999; 99US-0160981.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161982.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 96.6%; Score 28; DB 21; Length 1017;
 Best Local Similarity 83.3%; Pred. No. 1.3e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
 DB 93 EVRDLR 98

RESULT 11
 ABB72331
 ID ABB72331 standard; Protein; 265 AA.

AC ABB72331;

DT 04-APR-2002 (first entry)

DE Rat protein isolated from skin cells SEQ ID NO: 655.

KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
 KW developmental defect; inflammatory disease; dermatological; vulnery;
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.

OS Rattus sp.

PN WO200190357-A1.

PD 29-NOV-2001.

PF 24-MAY-2001; 2001MO-NZ00099.

PR 24-MAY-2000; 2000US-206550P.

PR 25-JUL-2000; 2000US-22132P.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Watson JD, Strachan L, Sleeman M, Onrust R, Murlison JG, Kumble KD;

DR WPI; 2002-122020/16.
 DR N-PSDB; ABL35016.

PT New polynucleotides and polypeptides encoded by the polynucleotides
 PT isolated from skin cells, useful for treating skin wounds, cancers,
 PT growth and developmental defects, inflammatory diseases, or for
 PT modulating immune responses

PS Claim 4; Page 411-412; 466pp; English.

CC The present invention provides the protein and coding sequences of cDNAs
 CC isolated from human, murine and rat skin cell libraries. The sequences
 CC can be used in the development of therapeutic agents useful in the
 CC treatment of skin diseases, including skin wounds, cancer, growth

CC defects, developmental defects and inflammatory diseases. The proteins
 CC have important roles in the induction of hair growth, cell proliferation
 CC and cell-cell interaction, in maintaining tissue integrity, in wound
 CC healing and in modulating immune responses. The present sequence is a
 CC polypeptide of the invention.

XX Sequence 265 AA;

Query Match 93.1%; Score 27; DB 23; Length 265;
 Best Local Similarity 83.3%; Pred. No. 5.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
 DB 132 ELRDLR 137

RESULT 12
 ABB49729
 ID ABB49729 standard; Protein; 299 AA.

AC ABB49729;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #2433.

KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.

XX Listeria monocytogenes.

PN WO200177335-A2.

PD 18-OCT-2001.

PF 11-APR-2001; 2001MO-FR01118.

PR 11-APR-2000; 2000FR-0004629.

PA (INSP) INST PASTEUR.

PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Faht H, Dehoux P,
 PI Dussurget O, Chetoui F, Nedjati H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Doman E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Madueno E, De Pablos B, Wehlend J, Kaerst U, Ertian K, Hauf J;
 PI Rose M, Voss H;

DR WPI; 2002-010914/01.

PT Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and
 PT related polypeptides

PS Claim 6; SEQ ID NO 2434; 192pp; French.

CC The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.
 CC monocytogenes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by L.

CC Monocytogenes and related organisms.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 299 AA;

Query Match 93.1%; Score 27; DB 23; Length 299;
 Best Local Similarity 83.3%; Pred. No. 5.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
 1:||||
 DB 181 EIRDLR 186

RESULT 13
 ID ABB59541
 ID ABB59541 standard; Protein; 323 AA.

AC ABB59541;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 5415.

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL03644.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 5415; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (ABBS737-ABBS72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 323 AA;

Query Match 93.1%; Score 27; DB 22; Length 323;
 Best Local Similarity 83.3%; Pred. No. 6.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
 1:||||
 DB 102 EIRDLR 107

DB 138 EIRDIR 143

RESULT 14

ID ABB27632
 ID ABB27632 standard; Protein; 554 AA.

AC ABB27632;

DT 02-JUL-2002 (first entry)

DE Streptococcus polypeptide SEQ ID NO 4440.

KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus agalactiae.

PN WO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-CB04789.

PR 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

PA (CHTR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;

DR WPI; 2002-352536/38.

DR N-PSDB; ABB68263.

PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -

PS Claim 1; Page 3605; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABB66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.

CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus (I) is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.

CC Sequence 554 AA;

Query Match 93.1%; Score 27; DB 23; Length 554;
 Best Local Similarity 83.3%; Pred. No. 1.1e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
 1:||||
 DB 102 EIRDIR 107

RESULT 15
 AAB31943
 ID AAB31943 standard; Protein: 555 AA.
 AC AAB31943;
 XX
 DT 15-MAY-2001 (first entry)
 DE Amino acid sequence of a partial dnaX (tau subunit) polypeptide.
 XX
 KW dnaE; Gram positive bacteria; polC; dnaE; holA, holB; dnaX; dnaN; ssb;
 KW dnaG; dnaB; antibiotic; replication; cell growth; cell death;
 KW bacterial infection.
 XX
 OS Streptococcus pyogenes.
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 148 /note= "encoded by TGT"
 FT Misc-difference 149 /note= "encoded by TGT"
 FT Misc-difference 150 /note= "encoded by TGT"
 FT Misc-difference 151 /note= "encoded by CTT"
 FT Misc-difference 152 /note= "encoded by TAT"
 FT Misc-difference 153 /note= "encoded by CTT"
 FT Misc-difference 154 /note= "encoded by GGC"
 FT Misc-difference 155 /note= "encoded by AAC"
 FT Misc-difference 156 /note= "encoded by AAC"
 FT Misc-difference 157 /note= "encoded by GGA"
 FT Misc-difference 157 /note= "encoded by ATG"
 XX
 PN WO200109164-A2.
 XX
 PD 08-FEB-2001.
 XX
 PF 28-JUL-2000; 2000WO-US20666.
 XX
 PR 29-JUL-1999; 99US-0146178.
 XX
 PA (UYRQ) UNIV ROCKEFELLER.
 XX
 PI O'Donnell ME, Bruck I, Zhang D, Whipple R;
 XX
 DR WPI: 2001-147453/15.
 DR N-PSDB; AAF54743.
 XX
 PT Isolated DNA molecule from a Gram positive bacterium encoding DNA
 PT replication proteins used to identify compounds which have antibiotic
 PT activity
 XX
 PS Claim 77; Page 60-61; 239pp; English.
 XX
 CC The present sequence represents a partial dnaX polypeptide. The
 CC specification describes DNA molecules from Gram positive bacteria,
 CC which comprise a coding region from a polC, dnaE, holA, holB, dnaX,
 CC dnaN, ssb, dnaG or a dnaB gene. These sequences encode proteins that
 CC replicate the chromosome of Gram positive bacteria. They are used for
 CC sequencing and amplification of DNA and in drug discovery to identify
 CC compounds which have antibiotic activity through interference with
 CC replication. They are used in methods for identifying compounds that
 CC are active at the level of DNA replication and result in arrest of
 CC cell growth or cell death of bacteria to treat bacterial infections in
 CC animals.
 CC
 XX
 SQ Sequence 555 AA;

Query Match 93.1%; Score 27; DB 22; Length 555;
 Best Local Similarity 83.3%; Pred. No. 1.1e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
 |||||:1
 DB 102 EIRDIR 107

Search completed: November 13, 2002, 13:22:11
 Job time : 30.7447 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 8.68085 Seconds
(without alignments)
20.336 Million cell updates/sec

Title: US-09-856-086-6

Perfect score: 29

Sequence: 1 EIRDLR 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	27	93.1	71 4 US-08-928-213B-24	Sequence 24, Appl
2	27	93.1	273 4 US-08-928-213B-11	Sequence 11, Appl
3	27	93.1	790 2 US-08-359-705B-9	Sequence 9, Appl
4	27	93.1	790 2 US-08-286-846A-9	Sequence 9, Appl
5	27	93.1	790 2 US-08-457-880A-9	Sequence 9, Appl
6	27	93.1	790 3 US-08-444-622A-9	Sequence 9, Appl
7	27	93.1	790 3 US-08-942-562-9	Sequence 9, Appl
8	27	93.1	790 4 US-09-156-923-9	Sequence 9, Appl
9	27	93.1	814 1 US-08-286-305A-3	Sequence 3, Appl
10	27	93.1	814 2 US-08-441-104A-3	Sequence 3, Appl
11	27	93.1	814 2 US-08-440-816A-3	Sequence 3, Appl
12	27	93.1	814 4 US-09-417-381A-3	Sequence 3, Appl
13	26	89.7	1184 4 US-09-541-782-2	Sequence 2, Appl
14	26	89.7	1184 4 US-09-723-820-2	Sequence 2, Appl
15	25	86.2	375 1 US-07-803-622E-7	Sequence 7, Appl
16	25	86.2	375 1 US-07-803-622E-9	Sequence 9, Appl
17	25	86.2	375 1 US-08-846-762-12	Sequence 12, Appl
18	25	86.2	420 3 US-08-685-871-58	Sequence 58, Appl
19	25	86.2	739 1 US-07-803-622E-2	Sequence 2, Appl
20	25	86.2	835 2 US-08-968-751-4	Sequence 4, Appl
21	25	86.2	1354 3 US-08-685-871-2	Sequence 2, Appl
22	24	82.8	78 4 US-09-134-001C-3279	Sequence 3279, Ap
23	24	82.8	134 4 US-09-134-001C-5232	Sequence 5232, Ap
24	24	82.8	200 4 US-08-949-155-4	Sequence 4, Appl
25	24	82.8	200 4 US-09-819-964-4	Sequence 4, Appl
26	24	82.8	207 2 US-08-177-109A-59	Sequence 59, Appl
27	24	82.8	207 2 US-08-687-706-59	Sequence 59, Appl

28	24	82.8	238 4 US-09-111-470-8	Sequence 8, Appl
29	24	82.8	309 4 US-09-347-803-14	Sequence 14, Appl
30	24	82.8	333 1 US-08-117-083-65	Sequence 65, Appl
31	24	82.8	344 4 US-09-134-001C-5004	Sequence 5004, Ap
32	24	82.8	348 4 US-08-855-910-8	Sequence 8, Appl
33	24	82.8	378 4 US-08-975-762-70	Sequence 70, Appl
34	24	82.8	378 4 US-09-295-028-70	Sequence 70, Appl
35	24	82.8	387 1 US-08-106-582-70	Sequence 70, Appl
36	24	82.8	387 1 US-08-713-828-3	Sequence 3, Appl
37	24	82.8	387 2 US-08-919-627-3	Sequence 3, Appl
38	24	82.8	387 2 US-09-096-245-3	Sequence 3, Appl
39	24	82.8	387 4 US-09-457-040B-30	Sequence 30, Appl
40	24	82.8	401 2 US-08-576-626A-45	Sequence 45, Appl
41	24	82.8	406 1 US-08-650-275-2	Sequence 2, Appl
42	24	82.8	406 3 US-09-181-318-2	Sequence 2, Appl
43	24	82.8	456 4 US-08-858-207A-366	Sequence 366, App
44	24	82.8	459 4 US-09-071-709-1	Sequence 1, Appl
45	24	82.8	526 2 US-08-898-780A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-928-213B-24
Sequence 24, Application US/08928213B
Patent No. 6238905
GENERAL INFORMATION:
APPLICANT: McHenry, Charles S.
Seville, Mark
Cull, Millard G.
TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III
NUMBER OF SEQUENCES: 195
CORRESPONDENCE ADDRESS:
ADDRESS: MEDLEN & CARROLL, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928, 213B
FILING DATE: 12-Sep-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MacKnight, Kamrin T.
REGISTRATION NUMBER: 38,230
REFERENCE/DOCKET NUMBER: ENZYCO-02550
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-705-8410
TELEFAX: 415-397-8338
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: Protein
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-08-928-213B-24
Query Match 93.1% Score 27; DB 4; Length 71;
Best Local Similarity 83.3%; Pred. No. 31;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EIRDLR 6
||||:|

Db 34 EIRDIR 39

RESULT 2
US-08-928-213B-11Sequence 11, Application US/08928213B
Patent No. 6238905

GENERAL INFORMATION:

APPLICANT: McHenry, Charles S.

Seville, Mark
Cull, Millard G.TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III
HOLENZYME

NUMBER OF SEQUENCES: 195

CORRESPONDENCE ADDRESSES:

ADDRESSEE: MEDLEN & CARROLL, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/928,213B

FILING DATE: 12-Sep-1997

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: MacKnight, Kamrin T.

REGISTRATION NUMBER: 38,230

REFERENCE/DOCKET NUMBER: ENZYCO-02550

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-705-8410

TELEFAX: 415-397-8338

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 273 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-08-928-213B-11

Query Match 93.1%; Score 27; DB 4; Length 273;

Best Local Similarity 83.3%; Pred. No. 1.3e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

Db 110 EIRDIR 115

RESULT 3
US-08-359-705B-9Sequence 9, Application US/08359705B
Patent No. 5844092

GENERAL INFORMATION:

APPLICANT: Presta, Leonard G.

APPLICANT: Shelton, David L.

APPLICANT: Ufer, Roman

TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/359,705B

FILING DATE: 20-Dec-1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/286846

FILING DATE: 08/10/94

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/215139

FILING DATE: 03/18/94

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, PhD., Timothy E.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: P0873P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/952-9881

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 790 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

US-08-359-705B-9

Query Match

Best Local Similarity 93.1%; Score 27; DB 2; Length 790;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

Db 83 EIRDLR 88

RESULT 4
US-08-286-846A-9Sequence 9, Application US/08286846A
Patent No. 5877016

GENERAL INFORMATION:

APPLICANT: Presta, Leonard G.

APPLICANT: Shelton, David L.

APPLICANT: Ufer, Roman

TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/286,846A

FILING DATE: 05-Aug-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, PhD., Timothy E.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: P0873P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-8674

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 790 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-286-846A-9

Query Match 93.1%; Score 27; DB 2; Length 790;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
|:||||
DB 83 EIRDLR 88

RESULT 5
US-08-457-880A-9
Sequence 9, Application US/08457880A
Patent No. 5910574
GENERAL INFORMATION:
APPLICANT: Leonard G. Presta
APPLICANT: David L. Shelton
APPLICANT: Roman Urfel
TITLE OF INVENTION: HUMAN TRK RECEPTORS AND NEUROTROPHIC FACTOR
TITLE OF INVENTION: INHIBITORS
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,880A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/444,622
FILING DATE: 19-May-1995
APPLICATION NUMBER: 08/286846
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0873P1C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 790 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-457-880A-9

Query Match 93.1%; Score 27; DB 2; Length 790;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
|:||||
DB 83 EIRDLR 88

RESULT 6
US-08-444-622A-9
Sequence 9, Application US/08444622A

Patent No. 6025166
GENERAL INFORMATION:
APPLICANT: Leonard G. Presta
APPLICANT: David L. Shelton
APPLICANT: Roman Urfel
TITLE OF INVENTION: HUMAN TRK RECEPTORS AND NEUROTROPHIC FACTOR INHIBITORS
TITLE OF INVENTION: INHIBITORS
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,622A
FILING DATE: 19-May-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286846
FILING DATE: 5
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0873P1C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 790 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-444-622A-9

Query Match 93.1%; Score 27; DB 3; Length 790;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
|:||||
DB 83 EIRDLR 88

RESULT 7
US-08-942-562-9
Sequence 9, Application US/08942562
Patent No. 6027927
GENERAL INFORMATION:
APPLICANT: Presta, Leonard G.
APPLICANT: Shelton, David L.
APPLICANT: Urfel, Roman
TITLE OF INVENTION: Human Trk Receptors and Neurotrophic
TITLE OF INVENTION: Factor Inhibitors
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/942,562
FILING DATE: 01-OCT-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/444,597
FILING DATE: 19-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, P.D., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P0873P1C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-8674
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 790 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-942-562-9

Query Match 93.1%; Score 27; DB 3; Length 790;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
1:|||||
DB 83 ELRLDR 88

RESULT 8
US-09-156-923-9
Sequence 9, Application US/09156923
Patent No. 6153189
GENERAL INFORMATION:
APPLICANT: Presta, Leonard G.
APPLICANT: Shelton, David L.
APPLICANT: Urie, Roman
TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/156,923
FILING DATE: 18-SEP-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/359,705
FILING DATE: 20-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/28646
FILING DATE: 10-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/215139
FILING DATE: 18-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Diegel, Ginger
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: GENENT.33CP2C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 949/760-0404
TELEFAX: 949/760-9502
INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 790 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-156-923-9

Query Match 93.1%; Score 27; DB 4; Length 790;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
1:|||||
DB 83 ELRLDR 88

RESULT 9
US-08-286-305A-3
Sequence 3, Application US/08286305A
Patent No. 5766863
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J.
APPLICANT: Mark, Melanie R.
APPLICANT: Sadlock, Michael D.
APPLICANT: Shelton, David L.
APPLICANT: Wong, Kai Lee Tan
TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,305A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/170558
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/157563
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 854C1P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 814 amino acids
TYPE: amino acid
TOPOLOGY: Linear
US-08-286-305A-3

Query Match 93.1%; Score 27; DB 1; Length 814;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
1:|||||
DB 107 ELRLDR 112

RESULT 10
US-08-441-104A-3
Sequence 3, Application US/08441104A
Patent No. 5891650
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J.
APPLICANT: Mark, Melanie R.
APPLICANT: Sadick, Michael D.
APPLICANT: Shelton, David L.
APPLICANT: Wong, Wai Lee Tan
TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,104A
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286305
FILING DATE: 05-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/170558
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/157563
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 854C1P1C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 814 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-441-104A-3
Query Match 93.1%; Score 27; DB 2; Length 814;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EIRDLR 6
|:||||
DB 107 ELRLDR 112

CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,816A
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/286305
FILING DATE: 05-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/170558
FILING DATE: 20-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/157563
FILING DATE: 23-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 854C1P1C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 814 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-440-816A-3
Query Match 93.1%; Score 27; DB 2; Length 814;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EIRDLR 6
|:||||
DB 107 ELRLDR 112

RESULT 12
US-09-417-381A-3
Sequence 3, Application US/09417381A
Patent No. 6287784
GENERAL INFORMATION:
APPLICANT: Godowski, Paul J.
APPLICANT: Mark, Melanie R.
APPLICANT: Sadick, Michael D.
APPLICANT: Shelton, David L.
APPLICANT: Wong, Wai Lee Tan
TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
FILE REFERENCE: P0854C1P2C1
CURRENT APPLICATION NUMBER: US/09/417,381A
CURRENT FILING DATE: 1999-10-13
NUMBER OF SEQ ID NOS: 11
SEQ ID NO 3
LENGTH: 814
TYPE: PRT
ORGANISM: Homo Sapien
US-09-417-381A-3
Query Match 93.1%; Score 27; DB 4; Length 814;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
DB 107 EIRDLR 112

RESULT 13
US-09-541-782-2
Sequence 2, Application US/09541782
Patent No. 6284480
GENERAL INFORMATION:
APPLICANT: Nislow, Corey
APPLICANT: Sakowicz, Roman
APPLICANT: Beraud, Christophe
TITLE OF INVENTION: Antifungal Assay
FILE REFERENCE: 1015
CURRENT APPLICATION NUMBER: US/09/541,782
CURRENT FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1184
TYPE: PRT
ORGANISM: A. nidulans
US-09-541-782-2

Query Match 89.7%; Score 26; DB 4; Length 1184;
Best Local Similarity 83.3%; Pred. No. 9.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
DB 680 EIRDLR 685

RESULT 14
US-09-723-820-2
Sequence 2, Application US/09723820
Patent No. 6468760
GENERAL INFORMATION:
APPLICANT: Nislow, Corey
APPLICANT: Sakowicz, Roman
APPLICANT: Beraud, Christophe
TITLE OF INVENTION: Antifungal Assay
FILE REFERENCE: 1015
CURRENT APPLICATION NUMBER: US/09/723,820
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 09/541,782
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1184
TYPE: PRT
ORGANISM: A. nidulans
US-09-723-820-2

Query Match 89.7%; Score 26; DB 4; Length 1184;
Best Local Similarity 83.3%; Pred. No. 9.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
DB 680 EIRDLR 685

RESULT 15
US-07-803-622E-7
Sequence 7, Application US/07803622E
Patent No. 5525497
GENERAL INFORMATION:
APPLICANT: Keller, Walter
APPLICANT: Lingner, Joachim

APPLICANT: Martin, Georges
APPLICANT: Wahle, Elmar
TITLE OF INVENTION: RECOMBINANT POLY(A) POLYMERASE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 611 West Sixth Street
CITY: Los Angeles
STATE: CA
COUNTRY: USA
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/803,622E
FILING DATE: 27-NOV-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 195/296
TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-489-1600
TELEFAX: 213-955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-803-622E-7

Query Match 86.2%; Score 25; DB 1; Length 375;
Best Local Similarity 66.7%; Pred. No. 4.8e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
DB 142 EIRDLR 147

Search completed: November 13, 2002, 13:18:09
Job time : 10.6809 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:18:18 : Search time 4.08511 Seconds
(without alignments)
22.121-Million cell updates/sec

Title: US-09-856-086-6
Sequence: 1 EIRDLR 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCIT_NEM_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEM_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEM_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCITUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEM_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEM_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEM_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	93.1	585	10	US-09-815-242-4973
2	27	93.1	585	10	US-09-815-242-10580
3	27	93.1	790	10	US-09-966-147-9
4	27	93.1	814	10	US-09-924-859A-3
5	26	89.7	87	10	US-09-764-878-151
6	26	89.7	325	10	US-09-741-669-945
7	26	89.7	501	10	US-09-815-242-5292
8	26	89.7	734	9	US-09-815-242-12503
9	26	89.7	734	9	US-10-008-355-5
10	25	86.2	177	10	US-09-815-242-5164
11	25	86.2	360	9	US-09-905-291A-213
12	25	86.2	360	10	US-09-925-289-861
13	25	86.2	360	10	US-09-909-330-213
14	25	86.2	360	10	US-09-909-088B-213
15	25	86.2	360	12	US-10-052-586-28
16	25	86.2	523	10	US-09-815-242-11918
17	25	86.2	663	10	US-09-815-242-11869
18	24	82.8	15	10	US-09-894-998-43
19	24	82.8	16	10	US-09-894-998-44

20	24	82.8	77	10	US-09-815-242-12214	Sequence 12214, A
21	24	82.8	77	10	US-09-815-242-12762	Sequence 12762, A
22	24	82.8	118	10	US-09-771-161A-97	Sequence 97, Appl
23	24	82.8	136	10	US-09-894-998-40	Sequence 40, Appl
24	24	82.8	139	10	US-09-864-761-44833	Sequence 44833, A
25	24	82.8	161	10	US-09-925-302-761	Sequence 761, App
26	24	82.8	223	10	US-09-925-300-1616	Sequence 1616, Ap
27	24	82.8	238	9	US-09-862-802-8	Sequence 8, Appl1
28	24	82.8	388	10	US-09-778-971-6	Sequence 6, Appl1
29	24	82.8	388	10	US-09-815-242-10718	Sequence 10718, A
30	24	82.8	376	10	US-09-894-998-39	Sequence 39, Appl
31	24	82.8	378	10	US-09-159-469-70	Sequence 70, Appl
32	24	82.8	378	10	US-09-798-042-70	Sequence 70, Appl
33	24	82.8	387	10	US-09-771-161A-188	Sequence 188, App
34	24	82.8	398	10	US-09-815-242-13171	Sequence 13171, A
35	24	82.8	406	9	US-09-978-295A-309	Sequence 309, App
36	24	82.8	406	10	US-09-765-205-8	Sequence 309, App
37	24	82.8	483	10	US-09-881-752A-354	Sequence 354, App
38	24	82.8	494	9	US-10-108-603-115	Sequence 115, App
39	24	82.8	502	10	US-09-886-404-18	Sequence 18, Appl
40	24	82.8	502	10	US-09-778-971-5	Sequence 5, Appl1
41	24	82.8	502	12	US-10-006-867-158	Sequence 158, App
42	24	82.8	502	12	US-10-052-586-400	Sequence 400, App
43	24	82.8	521	10	US-09-886-404-21	Sequence 21, Appl
44	24	82.8	524	10	US-09-764-853-731	Sequence 731, App
45	24	82.8	537	10	US-09-945-301-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-815-242-4973
Sequence 4973, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trivick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA, 011A
CURRENT FILING DATE: 2001-03-21
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: fastseq for Windows Version 4.0
SEQ ID NO 4973
LENGTH: 585
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-4973
Query Match 93.1%; Score 27; DB 10; Length 585;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;

Matches 5: Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
|||||
Db 103 EIRDIR 108

RESULT 2
US-09-815-242-10580

Sequence 10580, Application US/09815242
Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zykkind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes In

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10580

LENGTH: 586

TYPE: PRT

ORGANISM: Enterococcus faecalis

US-09-815-242-10580

Query Match

Best Local Similarity 93.1%; Score 27; DB 10; Length 586;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

|||||

Db 103 EIRDIR 108

RESULT 3

US-09-966-147-9

Sequence 9, Application US/09966147

Patent No. US20020146416A1

GENERAL INFORMATION:

APPLICANT: Presta, Leonard G.

APPLICANT: Shelton, David L.

APPLICANT: Ulfert, Roman

TITLE OF INVENTION: HUMAN TRK RECEPTORS AND NEUROTROPHIC FACTOR INHIBITORS

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear, LLP

STREET: 620 Newport Center Drive, 16th floor

CITY: Newport Beach

STATE: California

COUNTRY: USA

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 Inch, 1.44 MB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/966,147

FILING DATE: 27-Sep-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/446172

FILING DATE: 19-MAY-1995

APPLICATION NUMBER: 08/286846

FILING DATE: 05-AUG-1994

APPLICATION NUMBER: 08/215139

FILING DATE: 18-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Dreger, Ginger

REGISTRATION NUMBER: 33,055

REFERENCE/DOCKET NUMBER: GENEENT.33CPC4C

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 954-4114

TELEFAX: (415) 954-4111

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 790 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 9:

US-09-966-147-9

Query Match

Best Local Similarity 93.1%; Score 27; DB 10; Length 790;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

|||||

Db 83 EIRDLR 88

RESULT 4

US-09-924-859A-3

Sequence 3, Application US/09924859A

Patent No. US20020137113A1

GENERAL INFORMATION:

APPLICANT: Godowski, Paul J.

APPLICANT: Mark, Melanie R.

APPLICANT: Sadick, Michael D.

APPLICANT: Shelton, David L.

APPLICANT: Wong, Wai Lee Tan

TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY

FILE REFERENCE: P0854C1P2C1

CURRENT APPLICATION NUMBER: US/09/924,859A

PRIOR FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: 05/09/417,381

PRIOR FILING DATE: 1999-10-13

NUMBER OF SEQ ID NOS: 11

SEQ ID NO 3

LENGTH: 814

TYPE: PRT

ORGANISM: Homo Sapien

US-09-924-859A-3

Query Match

Best Local Similarity 93.1%; Score 27; DB 10; Length 814;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

|||||

Db 107 EIRDLR 112

RESULT 5

```
US-09-764-878-151
; Sequence 151, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA121
; CURRENT APPLICATION NUMBER: US/09/764,878
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 428
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 151
; LENGTH: 87
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-878-151

Query Match
Best Local Similarity 89.7%; Score 26; DB 10; Length 87;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 31 EIKDLR 36

RESULT 6
US-09-741-669-345
; Sequence 345, Application US/09741669
; Patent No. US20020022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Gene identification as required for
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 345
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-741-669-345

Query Match
Best Local Similarity 89.7%; Score 26; DB 10; Length 325;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 216 EIRDLR 221

RESULT 7
US-09-815-242-5292
; Sequence 5292, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Travick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
```

```
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5292
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(501)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-09-815-242-5292

Query Match
Best Local Similarity 89.7%; Score 26; DB 10; Length 501;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 134 DIRDRL 139

RESULT 8
US-09-815-242-12503
; Sequence 12503, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Travick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes In
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 12503
LENGTH: 504
TYPE: PRT
ORGANISM: *Staphylococcus aureus*
US-09-815-242-12503

Query Match 89.7%; Score 26; DB 10; Length 504;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
:|||||
DB 152 EIRDLR 157

RESULT 9
US-10-008-355-5
Sequence 5, Application US/10008355
Patent No. US20020164759A1
GENERAL INFORMATION:

APPLICANT: Travis, James
APPLICANT: Potempa, Jan S
APPLICANT: Banbula, Agnieszka
TITLE OF INVENTION: Dipeptidylpeptidases And Methods Of Use
FILE REFERENCE: 235.00440101
CURRENT APPLICATION NUMBER: US/10/008.355
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: US 60/246,827
PRIOR FILING DATE: 2000-11-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 734
TYPE: PRT
ORGANISM: *Shewanella putrefaciens*
US-10-008-355-5

Query Match 89.7%; Score 26; DB 9; Length 734;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
:|||||
DB 200 EIRDLR 205

RESULT 10
US-09-815-242-5164
Sequence 5164, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification Of Essential Genes In
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815.242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5164
LENGTH: 177
TYPE: PRT
ORGANISM: *Pseudomonas aeruginosa*
US-09-815-242-5164

Query Match 86.2%; Score 25; DB 10; Length 177;
Best Local Similarity 83.3%; Pred. No. 83;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EIRDLR 6
:|||||
DB 147 EIRDLR 152

RESULT 11
US-09-905-291A-213
Sequence 213, Application US/09905291A
Patent No. US20020160374A1
GENERAL INFORMATION:

APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Flaveroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905.291A
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15

PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 213
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens
US-09-905-291A-213

Query Match
Best Local Similarity 86.2%; Score 25; DB 9; Length 360;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EIRDLR 6
DB 132 EIRDLR 137

RESULT 12
US-09-925-299-861
Sequence 861, Application US/09925299
Patent No. US20020055627A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA102
CURRENT APPLICATION NUMBER: US/09/925,299
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05863
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1556
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 861
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (53)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (360)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-861

Query Match
Best Local Similarity 86.2%; Score 25; DB 10; Length 360;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 105 EYDVR 110

RESULT 13

US-09-909-320-213
Sequence 213, Application US/09909320
Patent No. US20020132240A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gueney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/909,320
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 213
LENGTH: 360
TYPE: PRT
ORGANISM: Homo sapiens

US-09-909-320-213

Query Match 86.2% Score 25; DB 10; Length 360;
Best Local Similarity 83.3% Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
||:||||
Db 132 EIQDLR 137

RESULT 14

US-09-909-088B-213
; Sequence 213, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrare, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gruney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/909, 088B
; CURRENT FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143, 048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145, 698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146, 222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095

; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 213
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-088B-213

Query Match 86.2% Score 25; DB 10; Length 360;
Best Local Similarity 83.3% Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
||:||||
Db 132 EIQDLR 137

RESULT 15

US-10-052-586-28
; Sequence 28, Application US/10052586
; Patent No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACTS ENCODING THE SAME
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052, 586
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120

; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089908

Query Match 86.2%; Score 25; DB 12; Length 360;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
Db 132 EIODLR 137

Search completed: November 13, 2002, 13:40:19
Job time : 4.22796 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:16:18 ; Search time 120.383 Seconds
(without alignments)
32.134 Million cell updates/sec

Title: US-09-856-086-6
Perfect score: 29
Sequence: 1 EIRDLR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1:	/cgn2_6/ptodata/1/paa/US06.COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06.COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07.COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US08.COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US081.COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US082.COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US083.COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US084.COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US085.COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US086.COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US087.COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US088.COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US089.COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US090.COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US091.COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US092.COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US093.COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US094.COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US095.COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US096.COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US097.COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US098.COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US099.COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US100.COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US101.COMB.pep.*
26:	/cgn2_6/ptodata/1/paa/US102.COMB.pep.*
27:	/cgn2_6/ptodata/1/paa/US60.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	122	21	US-09-791-537-21750
2	29	100.0	426	21	US-09-791-537-62039
3	29	100.0	542	21	US-09-791-537-92908
4	29	100.0	542	21	US-09-791-537-125497
5	29	100.0	542	26	US-10-205-331-57
6	29	100.0	543	20	US-09-611-526-3132

7	29	100.0	543	21	US-09-791-537-34475	Sequence 34475, A
8	29	100.0	543	21	US-09-791-537-56903	Sequence 56903, A
9	29	100.0	544	21	US-09-791-537-86802	Sequence 86802, A
10	29	100.0	548	21	US-09-791-537-125495	Sequence 125495, A
11	29	100.0	554	21	US-09-791-537-94465	Sequence 94465, A
12	29	100.0	555	21	US-09-791-537-18635	Sequence 18635, A
13	29	100.0	564	27	US-60-243-468-985	Sequence 985, App
14	29	100.0	768	25	US-10-104-047-3379	Sequence 3379, App
15	29	100.0	1161	16	US-09-253-991A-22872	Sequence 22872, A
16	29	100.0	53	20	US-09-621-976-5502	Sequence 5502, App
17	28	96.6	53	20	US-09-621-976-5506	Sequence 5506, App
18	28	96.6	53	27	US-60-147-459-5502	Sequence 5502, App
19	28	96.6	53	27	US-60-147-459-5506	Sequence 5506, App
20	28	96.6	118	21	US-09-708-427-15204	Sequence 15204, A
21	28	96.6	131	21	US-09-708-427-15203	Sequence 15203, A
22	28	96.6	133	21	US-09-708-427-15202	Sequence 15202, A
23	28	96.6	145	20	US-09-614-450-961	Sequence 961, App
24	28	96.6	220	21	US-09-791-537-38508	Sequence 38508, A
25	28	96.6	254	23	US-09-935-625-29828	Sequence 29828, A
26	28	96.6	254	27	US-60-361-742-1687	Sequence 1687, App
27	28	96.6	268	23	US-09-935-625-29066	Sequence 29066, A
28	28	96.6	338	19	US-09-513-996A-23086	Sequence 23086, A
29	28	96.6	338	19	US-09-513-996A-45820	Sequence 45820, A
30	28	96.6	338	21	US-09-791-537-147528	Sequence 147528, A
31	28	96.6	338	23	US-09-935-625-27644	Sequence 27644, A
32	28	96.6	338	23	US-09-935-625-29772	Sequence 29772, A
33	28	96.6	361	23	US-09-935-625-29831	Sequence 29831, A
34	28	96.6	556	27	US-60-360-039-4879	Sequence 4879, App
35	28	96.6	618	27	US-09-791-537-107216	Sequence 107216, App
36	28	96.6	618	27	US-60-360-039-1069	Sequence 1069, App
37	28	96.6	657	21	US-09-791-537-10864	Sequence 10864, App
38	28	96.6	752	27	US-60-361-742-1289	Sequence 1289, App
39	28	96.6	1017	19	US-09-513-996A-47640	Sequence 47640, A
40	28	96.6	1024	21	US-09-791-537-93333	Sequence 93333, A
41	28	96.6	1024	21	US-60-360-039-520	Sequence 520, App
42	27	93.1	71	17	US-09-151-888-24	Sequence 24, App
43	27	93.1	87	27	US-60-143-993-1873	Sequence 1873, App
44	27	93.1	92	16	US-09-252-991A-30642	Sequence 30642, A
45	27	93.1	117	15	US-09-134-000-5522	Sequence 5522, App

ALIGNMENTS

RESULT 1
US-09-791-537-21750
Sequence 21750, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biocomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 21750
LENGTH: 122
TYPE: PRT
ORGANISM: Mus musculus
US-09-791-537-21750

Query Match: 100.0%; Score 29; DB 21; Length 122;
Best Local Similarity: 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 50 EIRDLR 55


```
RESULT 2
US-09-791-537-62039
; Sequence 62039, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 62039
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Aeropyrum pernix
US-09-791-537-62039

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 426;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
DB 107 EIRDLR 112

RESULT 3
US-09-791-537-92908
; Sequence 92908, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 92908
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-92908

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 542;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
DB 142 EIRDLR 147

RESULT 4
US-09-791-537-125497
; Sequence 125497, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 125497
```

```
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-125497

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 542;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
DB 142 EIRDLR 147

RESULT 5
US-10-205-331-57
; Sequence 57, Application US/10205331
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alstair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pimock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205.331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: NF-L
US-10-205-331-57

Query Match
Best Local Similarity 100.0%; Score 29; DB 26; Length 542;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
DB 142 EIRDLR 147

RESULT 6
US-09-611-526-3132
; Sequence 3132, Application US/09611526
; GENERAL INFORMATION:
; APPLICANT: OYA, TOSHIO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: KAWAI, YURI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: KOJIMA, SHINICHI
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: KOGA, HISASHI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH CDNAS
; FILE REFERENCE: 08335/0122
; CURRENT APPLICATION NUMBER: US/09/611.526
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: JP 1999-194486
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: JP 2000-118774
; PRIOR FILING DATE: 2000-01-11
```

PRIOR APPLICATION NUMBER: JP 2000-183765
 PRIOR FILING DATE: 2000-05-02
 NUMBER OF SEQ ID NOS: 4484
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 3132
 LENGTH: 543
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-611-526-3132

Query Match 100.0%; Score 29; DB 20; Length 543;
 Best Local Similarity 100.0%; Pred. No. 2.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
 |||||
 DB 142 EIRDLR 147

RESULT 7
 US-09-791-537-34475
 Sequence 34475, Application US/09791537
 GENERAL INFORMATION:
 APPLICANT: Bionomix, Inc.
 APPLICANT: Dede, Derek
 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
 TITLE OF INVENTION: METHODS OF USE THEREOF
 FILE REFERENCE: 261/210
 CURRENT APPLICATION NUMBER: US/09/791,537
 CURRENT FILING DATE: 2001-02-22
 NUMBER OF SEQ ID NOS: 153055
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 34475
 LENGTH: 543
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-791-537-34475

Query Match 100.0%; Score 29; DB 21; Length 543;
 Best Local Similarity 100.0%; Pred. No. 2.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
 |||||
 DB 142 EIRDLR 147

RESULT 8
 US-09-791-537-56903
 Sequence 56903, Application US/09791537
 GENERAL INFORMATION:
 APPLICANT: Bionomix, Inc.
 APPLICANT: Dede, Derek
 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
 TITLE OF INVENTION: METHODS OF USE THEREOF
 FILE REFERENCE: 261/210
 CURRENT APPLICATION NUMBER: US/09/791,537
 CURRENT FILING DATE: 2001-02-22
 NUMBER OF SEQ ID NOS: 153055
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 56903
 LENGTH: 543
 TYPE: PRT
 ORGANISM: Mus musculus
 US-09-791-537-56903

Query Match 100.0%; Score 29; DB 21; Length 543;
 Best Local Similarity 100.0%; Pred. No. 2.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6

DB 142 EIRDLR 147
 |||||

RESULT 9
 US-09-791-537-86802
 Sequence 86802, Application US/09791537
 GENERAL INFORMATION:
 APPLICANT: Bionomix, Inc.
 APPLICANT: Dede, Derek
 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
 TITLE OF INVENTION: METHODS OF USE THEREOF
 FILE REFERENCE: 261/210
 CURRENT APPLICATION NUMBER: US/09/791,537
 CURRENT FILING DATE: 2001-02-22
 NUMBER OF SEQ ID NOS: 153055
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 86802
 LENGTH: 544
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-791-537-86802

Query Match 100.0%; Score 29; DB 21; Length 544;
 Best Local Similarity 100.0%; Pred. No. 2.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
 |||||
 DB 141 EIRDLR 146

RESULT 10
 US-09-791-537-125495
 Sequence 125495, Application US/09791537
 GENERAL INFORMATION:
 APPLICANT: Bionomix, Inc.
 APPLICANT: Dede, Derek
 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
 TITLE OF INVENTION: METHODS OF USE THEREOF
 FILE REFERENCE: 261/210
 CURRENT APPLICATION NUMBER: US/09/791,537
 CURRENT FILING DATE: 2001-02-22
 NUMBER OF SEQ ID NOS: 153055
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 125495
 LENGTH: 548
 TYPE: PRT
 ORGANISM: Sus scrofa
 US-09-791-537-125495

Query Match 100.0%; Score 29; DB 21; Length 548;
 Best Local Similarity 100.0%; Pred. No. 2.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
 |||||
 DB 140 EIRDLR 145

RESULT 11
 US-09-791-537-94465
 Sequence 94465, Application US/09791537
 GENERAL INFORMATION:
 APPLICANT: Bionomix, Inc.
 APPLICANT: Dede, Derek
 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
 TITLE OF INVENTION: METHODS OF USE THEREOF
 FILE REFERENCE: 261/210
 CURRENT APPLICATION NUMBER: US/09/791,537

QY 1 EIRDLR 6

CURRENT FILING DATE: 2001-07-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 94465
LENGTH: 554
TYPE: PRT
ORGANISM: Bos taurus
US-09-791-537-94465

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 554;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 141 EIRDLR 146

RESULT 12
US-09-791-537-18635
Sequence 18635, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 18635
LENGTH: 555
TYPE: PRT
ORGANISM: Bos taurus
US-09-791-537-18635

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 555;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 142 EIRDLR 147

RESULT 13
US-60-243-468-985
Sequence 985, Application US/60243468
GENERAL INFORMATION:
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
FILE REFERENCE: CL000929
CURRENT APPLICATION NUMBER: US/60/243,468
CURRENT FILING DATE: 2000-10-27
NUMBER OF SEQ ID NOS: 2121
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 985
LENGTH: 568
TYPE: PRT
ORGANISM: HUMAN
US-60-243-468-985

Query Match
Best Local Similarity 100.0%; Score 29; DB 27; Length 568;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 411 EIRDLR 416

RESULT 14
US-10-104-047-2379
Sequence 2379, Application US/10104047
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: NOVEL FULL LENGTH CDNA
FILE REFERENCE: H1-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2379
LENGTH: 764
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-2379

Query Match
Best Local Similarity 100.0%; Score 29; DB 25; Length 764;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 395 EIRDLR 400

RESULT 15
US-09-252-991A-22872
Sequence 22872, Application US/09252991A
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22872
LENGTH: 1161
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22872

Query Match
Best Local Similarity 100.0%; Score 29; DB 16; Length 1161;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 428 EIRDLR 433

Search completed: November 13, 2002, 13:39:20
Job time : 122.383 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:17:03 ; Search time 2.42553 Seconds
(Without alignments)
29.431 Million cell updates/sec

Title: US-09-856-086-6
Perfect score: 29
Sequence: 1 EIRDLR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41632 seqs, 11897606 residues

Total number of hits satisfying chosen parameters: 41632

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	6	5 US-09-856-086-6	Sequence 6, Appl1
2	29	100.0	8	5 US-09-856-086-7	Sequence 7, Appl1
3	25	86.2	360	6 US-10-125-923A-28	Sequence 28, Appl1
4	25	86.2	1120	6 US-10-240-851-95	Sequence 95, Appl1
5	24	82.8	78	6 US-10-092-411A-3279	Sequence 3279, Ap
6	24	82.8	116	1 PCT-US02-31996-16	Sequence 16, Appl1
7	24	82.8	134	6 US-10-092-411A-5232	Sequence 5232, Ap
8	24	82.8	236	1 PCT-US02-31996-12	Sequence 12, Appl1
9	24	82.8	282	1 PCT-US02-32851-4	Sequence 4, Appl1
10	24	82.8	319	6 US-10-264-213-149	Sequence 149, App
11	24	82.8	332	6 US-10-274-694-2	Sequence 2, Appl1
12	24	82.8	344	6 US-10-092-411A-5004	Sequence 5004, Ap
13	24	82.8	406	6 US-10-145-087A-309	Sequence 309, App
14	24	82.8	406	6 US-10-143-031A-309	Sequence 309, App
15	24	82.8	406	6 US-10-145-092A-309	Sequence 309, App
16	24	82.8	406	6 US-10-162-522A-309	Sequence 309, App
17	24	82.8	406	6 US-10-165-038A-309	Sequence 309, App
18	24	82.8	406	6 US-10-165-353-309	Sequence 309, App
19	24	82.8	406	6 US-10-170-481A-309	Sequence 309, App
20	24	82.8	406	6 US-10-172-039A-309	Sequence 309, App
21	24	82.8	406	6 US-10-145-016A-309	Sequence 309, App
22	24	82.8	406	6 US-10-145-088A-309	Sequence 309, App
23	24	82.8	406	6 US-10-145-139A-309	Sequence 309, App
24	24	82.8	406	6 US-10-131-813A-82	Sequence 82, Appl1
25	24	82.8	406	6 US-10-131-813A-430	Sequence 430, App
26	24	82.8	406	6 US-10-131-819A-82	Sequence 82, Appl1

27	24	82.8	406	6 US-10-131-819A-430	Sequence 430, App
28	24	82.8	406	6 US-10-131-823A-82	Sequence 82, Appl1
29	24	82.8	406	6 US-10-131-823A-430	Sequence 430, App
30	24	82.8	406	6 US-10-131-824A-82	Sequence 82, Appl1
31	24	82.8	406	6 US-10-131-824A-430	Sequence 430, App
32	24	82.8	406	6 US-10-131-826A-82	Sequence 82, Appl1
33	24	82.8	406	6 US-10-131-826A-430	Sequence 82, Appl1
34	24	82.8	406	6 US-10-131-829A-82	Sequence 82, Appl1
35	24	82.8	406	6 US-10-131-829A-430	Sequence 82, Appl1
36	24	82.8	406	6 US-10-125-926A-82	Sequence 82, Appl1
37	24	82.8	406	6 US-10-125-926A-430	Sequence 430, App
38	24	82.8	406	6 US-10-127-829A-82	Sequence 82, Appl1
39	24	82.8	406	6 US-10-127-829A-430	Sequence 82, Appl1
40	24	82.8	406	6 US-10-127-831A-82	Sequence 82, Appl1
41	24	82.8	406	6 US-10-127-831A-430	Sequence 82, Appl1
42	24	82.8	406	6 US-10-127-835A-82	Sequence 82, Appl1
43	24	82.8	406	6 US-10-127-835A-430	Sequence 82, Appl1
44	24	82.8	406	6 US-10-127-837A-82	Sequence 82, Appl1
45	24	82.8	406	6 US-10-127-837A-430	Sequence 430, App

ALIGNMENTS

RESULT 1
US-09-856-086-6
Sequence 6, Application US/09856086
GENERAL INFORMATION:
APPLICANT: EBRINGER, ALAN
TITLE OF INVENTION: DIAGNOSIS OF DEMYELINATING OR SPONGIFORM DISEASE
FILE REFERENCE: 78104,040
CURRENT APPLICATION NUMBER: US/09/856,086
CURRENT FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 6
TYPE: PRT
ORGANISM: Homo sapiens, Bos
US-09-856-086-6
Query Match
Best Local Similarity 100.0%; Score 29; DB 5; Length 6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EIRDLR 6
DB 1 EIRDLR 6
RESULT 2
US-09-856-086-7
Sequence 7, Application US/09856086
GENERAL INFORMATION:
APPLICANT: EBRINGER, ALAN
TITLE OF INVENTION: DIAGNOSIS OF DEMYELINATING OR SPONGIFORM DISEASE
FILE REFERENCE: 78104,040
CURRENT APPLICATION NUMBER: US/09/856,086
CURRENT FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 8
TYPE: PRT
ORGANISM: Homo sapiens, Bos
US-09-856-086-7
Query Match
Best Local Similarity 100.0%; Score 29; DB 5; Length 8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EIRDLR 6

DB 3 EIRDLR 8

RESULT 3
US-10-125-923A-28

; Sequence 28, Application US/10125923A

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey J.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C79

; CURRENT FILING DATE: 2002-01-15

; PRIOR FILING DATE: 2002-01-15

; PRIOR FILING DATE: 2002-01-15

; PRIOR FILING DATE: 1997-09-18

; PRIOR FILING DATE: 1997-09-18

; PRIOR FILING DATE: 1997-10-17

; PRIOR FILING DATE: 1997-10-24

; PRIOR FILING DATE: 1997-10-24

; PRIOR FILING DATE: 1997-10-24

; PRIOR FILING DATE: 1997-10-21

; PRIOR FILING DATE: 1997-10-21

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: US 09/229,319

; PRIOR FILING DATE: 1999-01-13

; PRIOR APPLICATION NUMBER: US 60/071,449

; PRIOR FILING DATE: 1998-01-13

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: US 60/105,511

PCT-US02-31996-16

Query Match 82.8%; Score 24; DB 1; Length 116;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 IRDLR 6
DB 34 IRDLR 38

RESULT 7

US-10-092-411A-5232
; Sequence 5232, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 5232
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-5232

Query Match 82.8%; Score 24; DB 6; Length 134;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRDLR 5
DB 127 IRDLR 131

RESULT 8

PCT-US02-31996-12
; Sequence 12, Application PC/TUS0231996
; GENERAL INFORMATION:
; APPLICANT: IMMUNEX CORPORATION
; APPLICANT: Butz, Eric A.
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: MAMMALIAN C-TYPE LECTINS
; FILE REFERENCE: 3318-NO
; CURRENT APPLICATION NUMBER: PCT/US02/31996
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: US 60/328,026
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Mus sp.
PCT-US02-31996-12

Query Match 82.8%; Score 24; DB 1; Length 236;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 IRDLR 6
DB 34 IRDLR 38

RESULT 9

PCT-US02-32851-4
; Sequence 4, Application PC/TUS0232851
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BECHA, Shanya D.
; APPLICANT: BHATIA, Umesh
; APPLICANT: BLAKE, Julie J.
; APPLICANT: BOROMSKY, Mark L.
; APPLICANT: BURRILL, John D.
; APPLICANT: CHANG, Hsin-Ru
; APPLICANT: CHAWLA, Nandinder K.
; APPLICANT: ELIOTT, Vicki S.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: FORSYTHE, Ian J.
; APPLICANT: GORVAD, Ann E.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: HO, Anne
; APPLICANT: ISON, Craig H.
; APPLICANT: KABLE, Amy E.
; APPLICANT: KHARE, Reena
; APPLICANT: LAL, Preeti G.
; APPLICANT: LEE, Sally
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LEE, Soo Yeun
; APPLICANT: LEHR-MASON, Patricia M.
; APPLICANT: LI, Joana X.
; APPLICANT: LINDQUIST, Erika A.
; APPLICANT: LUO, Wen
; APPLICANT: MARQUIS, Joseph P.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: SPRAGUE, William W.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: TANG, Y. Tom
; APPLICANT: WARREN, Bridget A.
; APPLICANT: YANG, Junming
; APPLICANT: YUE, Henry
; APPLICANT: ZEBARUDIAN, Yeganeh
; APPLICANT: ZHENG, Wenjin
; TITLE OF INVENTION: STRUCTURAL AND CYTOSKELETON-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1223 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/32851
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/328,931
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/360,681
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/343,896
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 60/346,308
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/332,385
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/340,776
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/347,703
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 7502092CD1
PCT-US02-32851-4

Query Match 82.8%; Score 24; DB 1; Length 282;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EIRDLR 6
Db 172 EARDLR 177

RESULT 10
US-10-264-213-149

Sequence 149, Application US/10264213
GENERAL INFORMATION:

APPLICANT: Glenn, Matthew
APPLICANT: Havukala, Ilkka J
APPLICANT: Lubbers, Mark William
TITLE OF INVENTION: Polynucleotides, materials incorporating
FILE REFERENCE: 11000.1043c3
CURRENT APPLICATION NUMBER: US/10/264,213
CURRENT FILING DATE: 2002-10-03
NUMBER OF SEQ ID NOS: 253
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 149
LENGTH: 319
TYPE: PRT
ORGANISM: Lactobacillus rhamnosus
US-10-264-213-149

Query Match 82.8%; Score 24; DB 6; Length 319;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
Db 71 KVRDLR 76

RESULT 11
US-10-274-694-2

Sequence 2, Application US/10274694
GENERAL INFORMATION:

APPLICANT: BAUGHN, Mariah R.
APPLICANT: BRUNS, Christopher M.
APPLICANT: DAS, Debopriya
DING, Li
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: GANDHI, Ameena R.
APPLICANT: HAFALIA, April J.A.
APPLICANT: KEARNEY, Liam
APPLICANT: KHAN, Farrah A.
APPLICANT: LAL, Preeti G.
APPLICANT: LEE, Ernestine A.
APPLICANT: LU, Dying Aina M.
APPLICANT: LU, Yan
APPLICANT: NGUYEN, Dannel B.
APPLICANT: PATTERSON, Chandra S.
APPLICANT: RAMKOWAR, Jayalaxmi
APPLICANT: RING, Huijun Z.
APPLICANT: SANJANWALA, Madhusudan M.
APPLICANT: TANG, Y. Tom
APPLICANT: THANGAVELU, Kavitha
APPLICANT: THORNTON, Michael B.
APPLICANT: TRIBOULET, Catherine M.
APPLICANT: WALIA, Narinder K.
APPLICANT: YANG, Junming
APPLICANT: YANG, Monique G.
APPLICANT: YUE, Henry
TITLE OF INVENTION: DRUG METABOLIZING ENZYMES
FILE REFERENCE: PI-0151 USA
CURRENT APPLICATION NUMBER: US/10/274,694
CURRENT FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: 60/221,837
PRIOR FILING DATE: 2000-07-28

PRIOR APPLICATION NUMBER: 60/220,037
PRIOR FILING DATE: 2000-07-21
PRIOR APPLICATION NUMBER: 60/218,948
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US01/21324
PRIOR FILING DATE: 2001-07-05
PRIOR APPLICATION NUMBER: 60/216,804
PRIOR FILING DATE: 2000-07-07
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PERL Program
SEQ ID NO 2
LENGTH: 332
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: 7477262CD1
US-10-274-694-2

Query Match 82.8%; Score 24; DB 6; Length 332;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
Db 239 KVRDLR 304

RESULT 12
US-10-092-411A-5004

Sequence 5004, Application US/10092411A
GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: 032796-101
CURRENT APPLICATION NUMBER: US/10/092,411A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5676
SEQ ID NO 5004
LENGTH: 344
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-092-411A-5004

Query Match 82.8%; Score 24; DB 6; Length 344;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 IRDLR 6
Db 17 IRDLR 21

RESULT 13
US-10-145-087A-309

Sequence 309, Application US/10145087A
GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang

```

: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630P1C47
: CURRENT APPLICATION NUMBER: US/10/145,087A
: PRIOR FILING DATE: 2001-10-18
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 309
: LENGTH: 406
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-145-087A-309

Query Match      82.8%; Score 24; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 EIRDL 5
Db      146 EIRDL 150

RESULT 14
US-10-143-031A-309
: Sequence 309, Application US/10143031A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Deenoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
```

```

: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630P1C39
: CURRENT APPLICATION NUMBER: US/10/143,031A
: PRIOR FILING DATE: 2002-10-10
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 624
: SEQ ID NO 309
: LENGTH: 406
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-143-031A-309

Query Match      82.8%; Score 24; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 EIRDL 5
Db      146 EIRDL 150

RESULT 15
US-10-145-092A-309
: Sequence 309, Application US/10145092A
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Deenoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
```


APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumes, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C45
CURRENT APPLICATION NUMBER: US/10/145,092A
CURRENT FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 624
SEQ ID NO 309
LENGTH: 406
TYPE: PRT
ORGANISM: Homo sapiens
US-10-145-092A-309

Query Match 82.8%; Score 24; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDL 5
Db 146 EIRDL 150

Search completed: November 13, 2002, 13:39:40
Job time : 3.42553 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2002, 13:13:25 ; Search time 9.31915 Seconds
(without alignments)
61.895 Million cell updates/sec

Title: US-09-856-086-6
Perfect score: 29
Sequence: 1 EIRDLR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 263224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 263224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	183	2	B64156
2	29	100.0	347	2	S44995
3	29	100.0	362	2	A44083
4	29	100.0	377	2	E84103
5	29	100.0	386	2	I39834
6	29	100.0	426	2	C72515
7	29	100.0	497	2	T51195
8	29	100.0	543	1	QFMSL
9	29	100.0	544	2	S07144
10	29	100.0	548	1	QFPGI
11	29	100.0	554	2	JW0094
12	29	100.0	958	2	E82994
13	28	96.6	133	2	B86238
14	28	96.6	220	2	G69192
15	28	96.6	285	1	S38567
16	28	96.6	347	2	D83653
17	28	96.6	430	1	VEKSGR
18	28	96.6	618	2	D64472
19	28	96.6	633	2	T28788
20	28	96.6	657	2	S05517
21	28	96.6	1024	2	F75489
22	28	96.6	1193	2	D82856
23	28	96.6	1864	1	UQ1657
24	27	93.1	299	2	AH1368
25	27	93.1	304	2	D75397
26	27	93.1	325	2	AZ2698
27	27	93.1	331	2	C72424
28	27	93.1	346	2	E97480
29	27	93.1	354	2	T39599

30	27	93.1	389	2	C96830	unknown protein F1
31	27	93.1	420	2	E75274	conserved hypotet
32	27	93.1	421	2	C83147	gamma-glutamyl pho
33	27	93.1	464	2	A56600	intermediate filam
34	27	93.1	528	2	S70295	probable membrane
35	27	93.1	563	2	S13786	DNA-directed DNA p
36	27	93.1	564	2	B83654	DNA polymerase III
37	27	93.1	579	2	AF1788	DNA polymerase III
38	27	93.1	579	2	AG1412	DNA polymerase III
39	27	93.1	693	2	T05006	hypothetical prote
40	27	93.1	790	1	TVHUTT	nerve growth facto
41	27	93.1	797	2	T33098	hypothetical prote
42	27	93.1	812	2	E75338	DNA gyrase, subun
43	27	93.1	848	2	T08858	vacuolar protein-s
44	27	93.1	914	2	T07065	probable lipoygen
45	26	89.7	88	2	E71896	hypothetical prote

ALIGNMENTS

RESULT 1

B64156
hypothetical protein HI0656 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #extl_change 19-May-2000
C:Accession: B64156
R:Flaeschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodex, A.; Kelley, J.M.; Weidman
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Georgagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:93530630; PMID:7542800
A:Accession: B64156
A:Molecule type: DNA
A:Status: nucleic acid sequence not shown; translation not shown
A:Residues: 1-183 <TIGR>
A:Cross-references: GB:U032748; GB:LA2023; NID:g1573645; PIDN:ANC22315.1; PID:g1573655
A:Note: best homolog was a hypothetical protein from Escherichia coli
C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MW0062

Query Match
Best Local Similarity 100.0%; Score 29; DB 2; Length 183;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
DB 170 EIRDLR 175

RESULT 2

S44995
pectate lyase - Erwinia carotovora
C:Species: Erwinia carotovora
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #extl_change 08-Oct-1999
C:Accession: S44995
R:Helkimo, R.; Flego, D.; Pirhonen, M.; Karlsson, M.B.; Eriksson, A.; Mae, A.; Ko
submitted to the EMBL Data Library, May 1994
A:Description: Characterization of a novel pectate lyase from Erwinia carotovora subs
A:Reference number: S44995
A:Accession: S44995
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-347 <HEI>
A:Cross-references: EMBL:X79232; NID:g488382; PIDN:CA55814.1; PID:g488383

Query Match
Best Local Similarity 100.0%; Score 29; DB 2; Length 347;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
|||||

DB 295 EIRDLR 300

RESULT 3

A4083

meg protein - Marek's disease virus

C:Species: Marek's disease virus

C>Date: 27-Jun-1994 #sequence_revision 27-Jun-1994 #text_change 01-Dec-1995

C:Accession: A44083

R:Jones, D.; Lee, L.; Liu, J.L.; Kung, H.J.; Tiliotson, J.K.

Proc. Natl. Acad. Sci. U.S.A. 89, 4042-4046, 1992

A:Title: Marek disease virus encodes a basic-leucine zipper gene resembling the fos/jun

A:Reference number: A44083; MUID:92237304; PMID:1315048

A:Accession: A44083

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-362 <JON>

A:Cross-references: GB:M89471

C:Superfamily: fos/jun DNA-binding domain homology

F:52-92/Domain: fos/jun DNA-binding domain homology <FJD>

Query Match 100.0%; Score 29; DB 2; Length 362;

Best Local Similarity 100.0%; Pred. No. 72;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

DB 102 EIRDLR 107

RESULT 4

E84103

two-component sensor histidine kinase involved in degradative enzyme degs [imported] - H

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: E84103

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20515582; PMID:11058132

A:Accession: E84103

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-377 <STO>

A:Cross-references: GB:AP001519; GB:BA000004; MUID:910176109; PIDN:BA07348.1; GSPDB:GNOC

A:Experimental source: strain C-125

C:Genetics:

A:Gene: degs

C:Superfamily: regulatory protein degs

Query Match 100.0%; Score 29; DB 2; Length 377;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

DB 217 EIRDLR 222

RESULT 5

I39834

protein kinase - Bacillus brevis

C:Species: Bacillus brevis

C>Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 26-Aug-1999

C:Accession: I39834

R:Low, M.E.; Reid, S.J.; James, M.D.; Watson, T.G.

Appl. Microbiol. Biotechnol. 42, 78-84, 1994

A:Title: Cloning and sequencing the degs-degU gene operon from an alkaliphilic Bacillus

A:Reference number: I39834; MUID:95169370; PMID:7765823

A:Accession: I39834

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-386 <RES>

A:Cross-references: GB:LI5444; MUID:9710494; PIDN:AA041438.1; PID:9710495.

C:Genetics:

A:Gene: degs

C:Superfamily: regulatory protein degs

Query Match 100.0%; Score 29; DB 2; Length 386;

Best Local Similarity 100.0%; Pred. No. 78;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

DB 226 EIRDLR 231

RESULT 6

C72515

hypothetical protein APE2096 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: C72515

R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.;

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki,

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Ae

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: C72515

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-426 <KAN>

A:Cross-references: DDBJ:AP000063; MUID:95105654; PIDN:BA081107.1; PID:95105795

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2096

C:Superfamily: conserved hypothetical protein Mui206

Query Match 100.0%; Score 29; DB 2; Length 426;

Best Local Similarity 100.0%; Pred. No. 87;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

DB 107 EIRDLR 112

RESULT 7

T51195

hypothetical protein B7N4.20 [imported] - Neurospora crassa

C:Species: Neurospora crassa

C>Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 28-Jul-2000

C:Accession: T51195

R:Schulte, U.; Algn, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakati

submitted to the Protein Sequence Database, July 2000

A:Reference number: 225286

A:Accession: T51195

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-497 <SCH>

A:Cross-references: EMBL:AL390218; GSPDB:GN00116; NCSP:B7N4.20

A:Experimental source: BAC clone B7N4; strain OR74A

C:Genetics:

A:Gene: NCSP:B7N4.20

A:Map position: 6

A:Introns: 342/1

Query Match 100.0%; Score 29; DB 2; Length 497;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

DB 211 EIRDLR 216

RESULT 8

GENSL

neurofilament triplet L protein - mouse

N:Alternate names: 68k neurofilament protein; NF-L(low) protein; type IV IF protein

C:Species: Mus musculus (house mouse)

C:Date: 31-Mar-1988 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999

C:Accession: A25227; A26562; A43772; A41012; I55316

R:Lewis, S.A.; Cowan, N.J.

MOL. CELL. BIOL. 6, 1529-1534, 1986

A:Title: Anomalous placement of introns in a member of the intermediate filament multigene

A:Reference number: A25227; MUID:87064433; PMID:3785173

A:Accession: A25227

A:Molecule type: DNA

A:Residues: 1-543 <LEW>

A:CROSS-references: GB:M3016; NID:9200023; PIDN:AAA39810.1; PID:9387492

A>Note: the authors translated the codon GGC for residue 5 as Ala, ACA for residue 88 as

1 as Glu

R:Lewis, S.A.; Cowan, N.J.

J. Cell Biol. 100, 843-850, 1985

A:Title: Genetics, evolution, and expression of the 68,000-mol-wt neurofilament protein:

A:Reference number: A26562; MUID:85131334; PMID:3919033

A:Accession: A26562

A:Molecule type: mRNA

A:Residues: 242-543 <LE2>

A:CROSS-references: GB:X02165

A:Experimental source: brain

R:Julien, J.P.; Meyer, D.; Flavel, D.; Hurst, J.; Grosfeld, F.

Brain Res. Mol. Brain Res. 1, 243-250, 1986

A:Title: Cloning and developmental expression of the murine neurofilament gene family.

A:Reference number: A43772

A:Accession: A43772

A:Molecule type: mRNA

A:Residues: 1-5, 'Y', '7-8, 'Y', '10-64, 'W', '66-72, 'L', '74-98, 'D', '100-194, 'R', '196-202, '204-239, 'X

A:CROSS-references: GB:M20480; NID:9200037; PIDN:AAA39814.1; PID:9200038

A>Note: the authors translated the codon CGC for residue 195 as Ala

R:Singh, R.K.; Nixon, R.A.

J. Biol. Chem. 266, 18861-18867, 1991

A:Title: Identification of Ser-55 as a major protein kinase A phosphorylation site on th

A:Reference number: A41012; MUID:92011653; PMID:1717455

A:Accession: A41012

A:Molecule type: protein

A:Residues: 52-57 <SR>

R:Nakamura, K.; Ikenaka, K.; Wada, K.; Tamura, T.

J. Biol. Chem. 265, 19786-19791, 1990

A:Title: Structure of the 68-kDa neurofilament gene and regulation of its expression.

A:Reference number: I55316; MUID:91060592; PMID:2246261

A:Accession: I55316

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-5, 'Y', '7-8, 'Y', '10-28 <RES>

A:CROSS-references: GB:M55423; NID:9200027; PIDN:AAA39812.1; PID:9554245

C:Comment: This is the most abundant of the three neurofilament proteins and, as the oth

C:Genetics:

A:Introns: 349/3; 391/2; 498/1

C:Superfamily: cytoskeletal keratin

C:Keywords: coiled coil; intermediate filament

DB 142 EIRDLR 147

RESULT 9

neurofilament triplet L protein - human

N:Contains: Glu-50 brain peptide

C:Species: Homo sapiens (man)

C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 21-Jul-2000

C:Accession: S07144; I52832; A60703

R:Julien, J.P.; Grosfeld, F.; Yazdabakhsh, K.; Flavel, D.; Melter, D.; Mushynski, W.

Biochem. Biophys. Acta 909, 10-20, 1987

A:Title: The structure of a human neurofilament gene (NF-L): a unique exon-intron org

A:Reference number: S07144; MUID:87214213; PMID:3034332

A:Accession: S07144

A:Molecule type: DNA

A:Residues: 1-544 <JUL>

A:CROSS-references: EMBL:X05608; NID:91495072; PIDN:CAA29097.1; PID:91279504

A>Note: the authors translated the codon ATG for residue 366 as Asn

R:Pospelov, V.A.; Pospelova, T.V.; Julien, J.P.

Cell Growth Differ. 5, 187-196, 1994

A:Title: Ap-1 and Krox-24 transcription factors activate the neurofilament light gene

A:Reference number: I52832; MUID:94235564; PMID:8180132

A:Accession: I52832

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-10 <POS>

A:CROSS-references: GB:S70309; NID:9547176; PIDN:AAI4057.1; PID:94261757

R:Nomata, Y.; Watanabe, T.; Wada, H.

J. Biochem. 93, 825-831, 1983

A:Title: Highly acidic proteins from human brain: purification and properties of Glu-

A:Reference number: A60703; MUID:83265667; PMID:6135695

A:Accession: A60703

A:Molecule type: protein

A:Residues: 469-472, 'D', '474 <NOM>

A:Experimental source: Glu-50 brain peptide

A>Note: this acidic protein is named for its greater than fifty per cent glutamic aci

C:Genetics:

A:Gene: GDB:NEFL; NFL

A:CROSS-references: GDB:120227; OMIM:162280

A:Map position: 8p21-8p21

A:Introns: 349/3; 391/2; 498/1

C:Superfamily: cytoskeletal keratin

C:Keywords: brain; coiled coil; intermediate filament

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

F:469-544/Product: Glu-50 peptide #status predicted <E50>

A:Title: Neurofilament architecture combines structural principles of intermediate filament
 A:Reference number: A90973
 A:Accession: A90973
 A:Molecule type: protein
 A:Residues: 1-83;278-548 <GE2>
 A:Note: residue 322 is either lysine or arginine
 R:Gonda, Y.; Nishikawa, K.; Ando, S.; Kitanura, S.; Minoura, Y.; Nishi, Y.; Inagaki, M.
 Biochem. Biophys. Res. Commun. 167, 1316-1325, 1990
 A:Title: Involvement of protein kinase C in the regulation of assembly-disassembly of neurofilament
 A:Reference number: A34569; MUID:90211318; PMID:2108674
 A:Accession: A34569
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 9-14;23-29;30-53 <GON>
 C:Comment: Mammalian neurofilaments usually contain three polypeptides, L, M, and H (with ke all other intermediate filament proteins) a conserved alpha-helical region, whose helical
 C:Comment: The amino-terminal headpiece is basic with a high content of hydroxylamino acid
 C:Comment: The extra mass and high charge density that distinguish the neurofilament protein
 C:Comment: The boundaries of the domains between residues 70-92 and 399-402 are not yet
 C:Comment: This protein was isolated from spinal cord.
 C:Keywords: coiled coil; intermediate filament
 F:1-70/Domain: head <HED>
 F:92-123/Domain: coil 1a, alpha-helical rod <RIA>
 F:137-232/Domain: coil 1b, alpha-helical rod <RIA>
 F:255-399/Domain: coil 2, alpha-helical rod <RIA>
 F:402-548/Domain: tail <TAI>
 F:403-442/Region: tail subdomain a
 F:443-548/Region: tail subdomain b

Query Match
 Best Local Similarity 100.0%; Score 29; DB 1; Length 548;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
 DB 140 EIRDLR 145

RESULT 11
 JMW094
 neurofilament protein-L - bovine
 N:Alternate names: NF-L
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 07-May-1999
 C:Accession: JMW094
 R:Hashimoto, R.; Nakamura, Y.; Goto, H.; Wada, Y.; Sakoda, S.; Kalbuchi, K.; Inagaki, M.
 Biochem. Biophys. Res. Commun. 245, 407-411, 1998
 A:Title: Domain- and site-specific phosphorylation of bovine NF-L by Rho-associated kinase
 A:Reference number: JMW094; MUID:98238650; PMID:9571164
 A:Accession: JMW094
 A:Molecule type: protein
 A:Residues: 1-554 <HAS>
 C:Comment: Domain- and site-specific phosphorylation by Rho-kinase regulates the assembly
 C:Superfamily: cytoskeletal keratin

Query Match
 Best Local Similarity 100.0%; Score 29; DB 2; Length 554;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
 DB 141 EIRDLR 146

RESULT 12
 E82994
 glycine cleavage system protein P1 PA5213 [Imported] - Pseudomonas aeruginosa (strain PA
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: E82994

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kae, A.; Lardig, K.;
 Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: E82994
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-958 <STO>
 A:Cross-references: GB:AE004334; GB:AE004091; NID:99951515; PIDN:AG08358.1; GSPDB:
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: gcvP1; PA5213

Query Match
 Best Local Similarity 100.0%; Score 29; DB 2; Length 958;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
 DB 225 EIRDLR 230

RESULT 13
 B86238
 protein T10024.3 [Imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: B86238
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Ali,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar
 ansen, N.F.; Hughes, B.; Huizart, L.
 Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Ki
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marz
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tai
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B86238
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-133 <STO>
 A:Cross-references: GB:AE005172; NID:95091534; PIDN:AD39563.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T10024.3
 A:Map position: 1

Query Match
 Best Local Similarity 96.6%; Score 28; DB 2; Length 133;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
 DB 19 EYRDLR 24

RESULT 14
 G69192
 ABC transporter (glutamine transport ATP-binding protein) - Methanobacterium thermo
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 02-Feb-2001
 C:Accession: G69192
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldege,
 i, Qiu, D.; Spadatore, R.; Vicalire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivani
 ki, S.; Church, G.M.; Daniels, C.D.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: f
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: G69192
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-220 <MTH>

A:Cross-references: GB:AE000849; GB:AE000666; NID:g2621780; PIDN:AAB85201.1; PID:g262178

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH696

C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop

F:21-213/Domain: ATP-binding cassette homology <ABC>

F:38-45/Region: nucleotide-binding motif A (P-loop)

Query Match

Best Local Similarity 96.6%; Score 28; DB 2; Length 220;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6

DB 5 EVRDLR 10

RESULT 15

S38567 phenylethanolamine N-methyltransferase (EC 2.1.1.28) - rat

N:Alternate names: noradrenaline N-methyltransferase

C:Species: Rattus norvegicus (Norway rat)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 05-May-2000

C:Accession: S38567; A60060; S03614; 156509

R:Sun, Y.H.; Kim, S.S.; Choi, W.; Hong, I.A.; Chong, Y.H.

Submitted to the EMBL Data Library, September 1993

A:Description: Structure and tissue specific expression of the rat phenylethanolamine N-

A:Reference number: S38567

A:Accession: S38567

A:Molecule type: DNA

A:Residues: 1-285 <STH>

A:Cross-references: EMBL:X75333; NID:g414186; PIDN:CA53082.1; PID:g414187

R:Welsberg, E.P.; Baruchin, A.; Stachowiak, M.K.; Stricker, E.M.; Zigmund, M.J.; Kaplan, B.

Brain Res. Mol. Brain Res. 6, 159-166, 1989

A:Title: Isolation of a rat adrenal cDNA clone encoding phenylethanolamine N-methyltrans

A:Reference number: A60060; M01D:90135920; PMID:2575695

A:Accession: A60060

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 'LA', 18-210, 'H', 212-213, 'L', 215-285 <WEI>

R:Mezey, E.

Nucleic Acids Res. 17, 2125, 1989

A:Title: Cloning of the rat adrenal medullary phenylethanolamine-N-methyltransferase.

A:Reference number: S03614; M01D:89183620; PMID:2928117

A:Accession: S03614

A:Molecule type: mRNA

A:Residues: 26-51, 'H', 53-203, 'A', 205-210, 'H', 212-213, 'L', 215-285 <MEZ>

A:Cross-references: EMBL:X14211; NID:g56943; PIDN:CA32428.1; PID:g56944

R:Sun, Y.H.; Chun, Y.S.; Lee, I.S.; Kim, S.S.; Choi, W.; Chong, Y.H.; Hong, L.; Kim, S.H.

J. Neurochem. 63, 1603-1608, 1994

A:Title: Complete nucleotide sequence and tissue-specific expression of the rat phenyleth

A:Reference number: 156509; M01D:95016700; PMID:7931317

A:Accession: 156509

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-285 <RES>

A:Cross-references: EMBL:X75333; NID:g414186; PIDN:CA53082.1; PID:g414187

C:Comment: This enzyme converts noradrenaline into adrenaline.

C:Comment: Increased expression of this protein occurs in the adrenal medulla of cold-st

C:Genetics: 69/1; 138/2

C:Superfamily: phenylethanolamine N-methyltransferase

C:Keywords: adrenal gland; methyltransferase; S-adenosylmethionine

Query Match

Best Local Similarity 96.6%; Score 28; DB 1; Length 285;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6

DB 5 EVRDLR 10

DB 250 EVRDLR 255

Search completed: November 13, 2002, 13:23:30
Job time : 11.462 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2002, 13:13:24 ; Search time 4.97872 Seconds
(without alignments)
49.984 Million cell updates/sec

Title: US-09-856-086-6
Perfect score: 29
Sequence: 1 EIRDLR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	183	YRDC_HAELIN	P44807 haemophilus
2	29	100.0	386	DEGS_BACBR	P54663 bacillus br
3	29	100.0	426	YK96_AERPE	O9YA42 aeropyrum p
4	29	100.0	541	NFL_RAT	P19527 rattus norv
5	29	100.0	542	NFL_MOUSE	P08551 mus musculu
6	29	100.0	543	NFL_HUMAN	P07196 homo sapien
7	29	100.0	548	NFL_PIG	P03547 sus scrofa
8	29	100.0	554	NFL_BOVIN	P03548 bos taurus
9	28	96.6	285	PNMT_RAT	P10937 rattus norv
10	28	96.6	430	GFAP_MOUSE	P03995 mus musculu
11	28	96.6	618	CAB2_METJA	O58776 methanococc
12	28	96.6	657	LAMA_CHICK	P13648 gallus gall
13	28	96.6	1024	CANB_DEIRA	O94WKO deiniococcu
14	28	96.6	1864	VCNB_RCMV	P33930 red clover
15	27	93.1	421	PROA_PSEAE	O9HX20 pseudomonas
16	27	93.1	464	IFE_BRALA	O04948 brachylostos
17	27	93.1	499	AIRX_HUMAN	O16352 homo sapien
18	27	93.1	528	YAC8_YEAST	P39734 saccharomyc
19	27	93.1	563	DPX3_BACSU	P03122 bacillus su
20	27	93.1	790	KIF9_MOUSE	O9WV04 mus musculu
21	27	93.1	796	TRKA_HUMAN	O64629 homo sapien
22	26	89.7	136	VG46_BPMO2	O64237 mycobacteri
23	26	89.7	193	E631_DROME	P48593 drosophilla
24	26	89.7	199	YF24_RTYRP	P27500 rice tungro
25	26	89.7	305	MPF8_HUMAN	O9S549 homo sapien
26	26	89.7	305	YBKR_ECOLI	P77367 escherichia
27	26	89.7	461	VIME_ONCMY	P46674 oncorhynchu
28	26	89.7	469	TRPC_PASMY	P57855 pasteurrella
29	26	89.7	475	KPKK_CORGL	O46078 corynebacte
30	26	89.7	640	PRCK_HUMAN	O16822 homo sapien
31	26	89.7	708	PRSD_MYXXA	P17531 myxococcus
32	26	89.7	739	YGIQ_ECOLI	O46861 escherichia
33	26	89.7	857	CAMP_THES7	P51060 thermus sp.

34	26	89.7	1163	1	SBCQ_CLOAB	O97fk1 clostridium
35	26	89.7	1184	1	BIMC_EMENT	P17120 ementisella
36	26	89.7	4499	1	DYHA_CHLRE	O39610 chlamydomon
37	25	86.2	111	1	DSR2_VIBCH	O85264 vibrio chol
38	25	86.2	112	1	INSN_SHIDY	P39213 shigella dy
39	25	86.2	134	1	INR1_ECOLI	P75679 escherichia
40	25	86.2	171	1	Y162_METH	O27890 methanobact
41	25	86.2	291	1	PYRB_AQUAE	O66726 aquilifex aeo
42	25	86.2	300	1	Y505_MYCLE	O49823 mycobacteri
43	25	86.2	308	1	Y505_MYCTU	O11169 mycobacteri
44	25	86.2	310	1	URB1_RAT	P51593 rattus norv
45	25	86.2	344	1	GCH2_PTCGU	P50139 pichia gull

ALIGNMENTS

RESULT 1
YRDC_HAELIN STANDARD; PRT; 183 AA.
AC P44807:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein HI0656.
GN HI0656.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
RN [2]
RP IDENTIFICATION BY MASS SPECTROMETRY.
RX MEDLINE=20137488; PubMed=10675023;
RA Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B.,
RA Gray C., Fountoulakis M.;
RL "Two-dimensional map of the proteome of Haemophilus influenzae.";
RL Electrophoresis 21:411-429(2000).
-1- FUNCTION: Binds preferentially to double-stranded RNA (By
similarity).
-1- SIMILARITY: BELONGS TO THE SNA5/YRDC/YCIO/WLC FAMILY. STRONG, TO
E. COLI YRDC.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL: U32748; AAC2315.1; -
HSSP: P45748; 1HRU.
TIGR: HI0656; -
InterPro: IPR000666; Sna5_ycio_yrdc.
Pfam: PF01300; Sna5_ycio_yrdc; 1.
ProSite: PS01147; SNA5; 1.
RNA-binding; Complete proteome.

```

SQ SEQUENCE 183 AA; 20487 MW; 7B52523B5341D71D CRC64;
Query Match 100.0%; Score 29; DB 1; Length 183;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
    |||||
DB 170 EIRDLR 175

RESULT 2
DEGS_BACBR STANDARD; PRT; 386 AA.
AC P54663;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sensor protein degs (EC 2.7.3.-).
GN DEGS.
OS Bacillus brevis.
OC Bacteria; Firmicutes; Bacillales; Paenibacillaceae; Brevibacillus.
OX NCBI_TaxID=1393;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ALK36.
RX MEDLINE=95169370; PubMed=7765823;
RA Louw M.E., Reid S.J., James M.D., Watson T.G.;
RT "Cloning and sequencing the degs-degu operon from an alkalophilic
RT Bacillus brevis.";
RL Appl. Microbiol. Biotechnol. 42:78-84(1994).
CC -1- FUNCTION: INVOLVED IN A SENSOR TRANSDUCTION PATHWAY THAT AFFECT
CC THE PRODUCTION OF ENZYMES THAT DEGRADE POLYMERIC CARBON AND
CC NITROGEN SOURCES. DEGS PROBABLY ACTS AS A KINASE THAT
CC PHOSPHORYLATES DEGU.
CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L15444; AAC41438.1;
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR004359; HIS_KIN_s1g.
DR Pfam: PF02518; HATPase_C; 1.
DR SMART: SM00387; HATPase_C; 1.
DR PROSITE: PS50109; HIS_KIN; 1.
KW Sensory transduction; Transferrase; Kinase; Phosphorylation.
FT MOD_RES 188 384 HISTIDINE KINASE.
FT DOMAIN 194 194 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 386 AA; 44837 MW; EE209CB8F663A17E CRC64;

Query Match 100.0%; Score 29; DB 1; Length 386;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
    |||||
DB 226 EIRDLR 231

RESULT 3
YK96_AERPE STANDARD; PRT; 426 AA.
ID YK96_AERPE
AC O9YA42;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein APE2096.

GN APE2096.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatakeya Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankel A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakatani H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
CC -1- SIMILARITY: BELONGS TO THE UPF0095 FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000063; BA81107.1;
DR InterPro: IPR002936; DNAPrim_toprim.
DR Pfam: PF01751; TOPrim; 1.
DR SMART: SM00493; TOPRIM; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 426 AA; 47681 MW; 73A0B78FCDD2106B1 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 426;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
    |||||
DB 107 EIRDLR 112

RESULT 4
NFL_RAT STANDARD; PRT; 541 AA.
ID NFL_RAT
AC P19527; O63367;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NFL-L).
GN NFL OR NFL OR NF68.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90184052; PubMed=2516804;
RA Chin S.S., Liem R.K.H.;
RT "Expression of rat neurofilament proteins NF-L and NF-M in
RT transfected non-neuronal cells.";
RL Eur. J. Cell Biol. 50:475-490(1989).
RN [2]
RP SEQUENCE OF 197-483 FROM N.A.
RX MEDLINE=85252830; PubMed=392599;
RA Julien J.-P., Ramchandran K., Grosved F.;
RT "Cloning of a cDNA encoding the smallest neurofilament protein from
RT the rat.";
RL Biochim. Biophys. Acta 825:398-404(1985).
RN [3]

```



```

RP SEQUENCE OF 1-10 FROM N.A.
RA MEDLINE:95264348; PubMed-7745611;
RX Reeben M., Neuman T., Palgi J., Palm K., Paalme V., Saarna M.;
RT "Characterization of the rat light neurofilament (NF-L) gene promoter
RT and identification of NGF and cAMP responsive regions.";
RL J. Neurosci. Res. 40:177-188(1995).
[4]
RP CARBOHYDRATE-LINKAGE SITES
RX MEDLINE-93346421; PubMed-8344946;
RA Dong D.L.-Y., Xu Z.-S., Chevrier M.R., Cotter R.J., Cleveland D.W.,
RA Hart G.W.;
RT "Glycosylation of mammalian neurofilaments. Localization of multiple
RT O-linked N-acetylglucosamine moieties on neurofilament polypeptides
RL L and M.";
RL J. Biol. Chem. 268:16679-16687(1993).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHANGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- PTM: O-GLYCOSYLATED; CONTAINS THREE N-ACETYLGLUCOSAMINE SIDE
CC CHAINS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONDEUTHERIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, AF031880; AAB87069.1; -
DR EMBL, M25638; AAA41694.1; -
DR EMBL, X53981; CAA37931.1; -
DR PIR: A21762; A21762.
DR GlycoSuiteDB: p19527; -
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
DR Intermediate filament; Colled coil; Neurone; Glycoprotein.
KM INT1MET 0
FT 0
FT DOMAIN 1 92 HEAD.
FT DOMAIN 93 396 ROD.
FT DOMAIN 397 541 TAIL.
FT DOMAIN 93 124 COIL 1A.
FT DOMAIN 125 137 LINKER 1.
FT DOMAIN 138 233 COIL 1B.
FT DOMAIN 234 252 LINKER 12.
FT DOMAIN 253 271 COIL 2A.
FT DOMAIN 272 280 LINKER 2.
FT DOMAIN 281 396 COIL 2B.
FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
FT DOMAIN 444 541 TAIL, SUBDOMAIN B (ACIDIC).
FT CARBOHD 20 20 O-LINKED (GLICNAC).
FT CARBOHD 26 26 O-LINKED (GLICNAC).
FT CARBOHD 391 391 /FTID-CAR.000129.
FT SITE 381 391 EPTIPE (RECOGNIZED BY IF-SPECIFIC
FT MONOCLONAL ANTIBODY).
FT CONFLICT 197 202 GAD6A -> KARMS (IN REF. 2).
FT CONFLICT 399 399 R -> K (IN REF. 2).
FT CONFLICT 476 476 A -> E (IN REF. 2).
FT CONFLICT 480 483 EKER -> KRDE (IN REF. 2).
SQ SEQUENCE 541 AA; 61204 MM; 0D17839A226918A CXC64;
Query Match 100.0%; Score 29; DB 1; Length 541;

```

```

Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 EIRDLR 6
          |||||
Db       141 EIRDLR 146

RESULT 5
NFL_MOUSE
ID NFL_MOUSE      STANDARD:      PRT: 542 AA.
AC P08551;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
DE NFL GN NFL.
OS Mus musculus (Mouse).
OC Eumariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OC NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=87064433; PubMed=3785173;
RA Lewis S.A., Cowan N.J.;
RT "Anomalous placement of introns in a member of the intermediate
RT filament multigene family: an evolutionary conundrum.";
RL Mol. Cell. Biol. 6:1529-1534(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=87158637; PubMed=3103856;
RA Julien J.-P., Meyer D., Flavell D., Hurst J., Grosved F.;
RT "Cloning and developmental expression of the murine neurofilament
RT gene family.";
RL Brain Res. 387:243-250(1986).
RN [3]
RP SEQUENCE OF 241-542 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=85131334; PubMed=3919033;
RA Lewis S.A., Cowan N.J.;
RT "Genetics, evolution, and expression of the 68,000-mol-wt
RT neurofilament protein: isolation of a cloned cDNA probe.";
RL J. Cell Biol. 100:843-850(1985).
RN [4]
RP SEQUENCE OF 1-27 FROM N.A.
RX MEDLINE=91060592; PubMed=2246261;
RA Nakahita K., Ikenaka K., Wada K., Tamura T.A., Furutachi T.,
RA Mikoshiba K.;
RT "Structure of the 68-kDa neurofilament gene and regulation of its
RT expression.";
RL J. Biol. Chem. 265:19786-19791(1990).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/)

```

```

CC      or send an email to license@slb.ch).
CC      -----
CC      EMBL: X02165; CAB51616.1; -
CC      EMBL: M20480; AAA39814.1; -
CC      EMBL: M13016; AAA39810.1; -
CC      EMBL: M55423; AAA39812.1; -
CC      PIR: A25227; CEMSL.
CC      MGD: MGI:97313; NFL.
CC      InterPro: IPR001664; IF.
CC      Pfam: PF00038; filament; 1.
CC      PROSITE: PS00226; IF; 1.
CC      Intermediate filament; Coiled coil; Neurone; Glycoprotein.
CC      INTR_MET 0 0
CC      DOMAIN 1 92 HEAD.
CC      DOMAIN 93 396 ROD.
CC      DOMAIN 397 542 TAIL.
CC      DOMAIN 93 124 COIL 1A.
CC      DOMAIN 125 137 LINKER 1.
CC      DOMAIN 138 233 COIL 1B.
CC      DOMAIN 234 252 LINKER 12.
CC      DOMAIN 253 271 COIL 2A.
CC      DOMAIN 272 280 LINKER 2.
CC      DOMAIN 281 396 COIL 2B.
CC      DOMAIN 397 443 TAIL, SUBDOMAIN A.
CC      DOMAIN 444 542 TAIL, SUBDOMAIN B (ACIDIC).
CC      CARBOHYD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
CC      CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
CC      SITE 381 391 EPTOPE (RECOGNIZED BY IF-SPECIFIC
CC      MONOCLONAL ANTIBODY).
CC      FT CONFLICT 5 5 Y -> S (IN REF. 1).
CC      FT CONFLICT 8 8 Y -> I (IN REF. 1).
CC      FT CONFLICT 64 64 M -> K (IN REF. 1).
CC      FT CONFLICT 72 72 V -> L (IN REF. 2).
CC      FT CONFLICT 98 98 D -> H (IN REF. 1).
CC      FT CONFLICT 194 194 R -> A (IN REF. 1).
CC      FT CONFLICT 202 202 MISSING (IN REF. 2).
CC      FT CONFLICT 239 239 Y -> I (IN REF. 1).
CC      SO SEQUENCE 542 AA; 61448 MW; 8EEB86CF0831D8C CRC64;

Query Match      100.0%; Score 29; DB 1; Length 542;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 EIRDLR 6
Db      141 EIRDLR 146

RESULT 6
NFL_HUMAN      STANDARD; PRT; 543 AA.
ID P07196; O16154;
AC P07196; O16154;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
GN NEFL OR NFL OR NF68.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
PI [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87214213; Pubmed=30343332;
RA Julien J.-P., Grosfeld F., Yazdankbakhsh K., Flavell D., Melzer D.,
RA Mshynski W.;
RT "The structure of a human neurofilament gene (NF-L): a unique exon-
RT Intron organization in the intermediate filament gene family.";
RL Biochim. Biophys. Acta 909:10-20(1987).
RN [2]
RP SEQUENCE OF 1-10 FROM N.A.
RX MEDLINE=94235564; Pubmed=8180132;

```

DB 140 EIRDLR 145

|||||

RESULT 7
ID NFL_PIG STANDARD; PRT; 548 AA.
AC P02547;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
GN NEFL.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=85154583; PubMed=3920075;
RA Geisler N., Plessmann U., Weber K.;
RT "The complete amino acid sequence of the major mammalian
RT neurofilament protein (NF-L)."
RL FEBS Lett. 182:475-478(1985).
RN [2]
RP SEQUENCE OF 1-82 AND 278-548.
RA Geisler N., Kaufmann E., Fischer S., Plessmann U., Weber K.;
RT "Neurofilament architecture combines structural principles of
RT intermediate filaments with carboxy-terminal extensions increasing
RT in size between triplet proteins."
RL EMBL J. 2:1295-1302(1983).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC PIR: A02963; ORFGL.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neurone; Glycoprotein.
FT DOMAIN 1 91 HEAD.
FT DOMAIN 92 395 ROD.
FT DOMAIN 396 548 TAIL.
FT DOMAIN 548 548 TAIL.
FT DOMAIN 92 123 COIL 1A.
FT DOMAIN 124 136 LINKER 1.
FT DOMAIN 137 232 COIL 1B.
FT DOMAIN 233 251 LINKER 12.
FT DOMAIN 252 270 COIL 2A.
FT DOMAIN 271 279 LINKER 2.
FT DOMAIN 280 395 TAIL.
FT DOMAIN 396 442 TAIL, SUBDOMAIN A.
FT DOMAIN 443 548 TAIL, SUBDOMAIN B (ACIDIC).
FT CARBOHYD 20 26 O-LINKED (GLCNAC) (BY SIMILARITY).
FT CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
FT SITE 380 390 EPTIDE (RECOGNIZED BY IF-SPECIFIC
FT MONOCLONAL ANTIBODY).
FT UNSURE 322 322 OR K.
SO SEQUENCE 548 AA; 61940 MW; 83044813637ACT739 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 548;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6

DB 140 EIRDLR 145

|||||

RESULT 8
ID NFL_BOVIN STANDARD; PRT; 554 AA.
AC P02548; P79127;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L) (Micro glutamic acid-rich
DE protein).
GN NEFL.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Brain;
RA Hill W.D., Zhang U., Balin B.J., Sprinkle T.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 468-554.
RX MEDLINE=85154567; PubMed=3884373;
RA Isobe T., Okuyama T.;
RT "Brain micro glutamic acid-rich protein is the C-terminal endpiece of
RT the neurofilament 68-kDa protein as determined by the primary
RT sequence."
RL FEBS Lett. 182:389-392(1985).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U83919; AAB41543.1;
CC PIR: A02964; OFBO.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neurone.
FT INIT_MET 0 0 BY SIMILARITY.
FT DOMAIN 1 92 HEAD (BY SIMILARITY).
FT DOMAIN 93 396 ROD (BY SIMILARITY).
FT DOMAIN 397 554 TAIL (BY SIMILARITY).
FT DOMAIN 554 554 TAIL.
FT DOMAIN 93 124 LINKER 1.
FT DOMAIN 125 137 LINKER 1.
FT DOMAIN 138 233 COIL 1B.
FT DOMAIN 234 252 LINKER 12.
FT DOMAIN 253 271 COIL 2A.
FT DOMAIN 272 280 LINKER 2.
FT DOMAIN 281 396 COIL 2B.
FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
FT DOMAIN 444 554 TAIL, SUBDOMAIN B (ACIDIC).

FT CONFLICT 494 500 MISSING (IN REF. 2).
 FT CONFLICT 509 509 A -> AEA (IN REF. 2).
 SO SEQUENCE 554 AA; 62514 MW; D772B81CA2C31CIA CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 554;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EIRDLR 6
 DB 141 EIRDLR 146
 RESULT 9
 PNMNT_RAT STANDARD; PRT; 285 AA.
 ID PNMNT_RAT
 AC P10937;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phenylethanolamine N-methyltransferase (EC 2.1.1.28) (PNMTase)
 DE (Norepinephrine N-methyltransferase).
 GN PNMT.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=95016700; PubMed=7931317;
 RA Suh Y.H., Chun Y.S., Lee I.S., Kim S.S., Choi W., Chong Y.H.,
 RA Hong L., Kim S.H., Park C.W., Kim C.G.;
 RT "Complete nucleotide sequence and tissue-specific expression of the
 RT rat phenylethanolamine N-methyltransferase gene.";
 RT J. Neurochem. 63:1603-1608(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mistar; TISSUE=Spleen;
 RX MEDLINE=96035937; PubMed=7558218;
 RA Kolke G., Jacob H.J., Krieger J.E., Szpirer C., Hoehne M.R.,
 RA Horluchl M., Deau V.J.;
 RT "Investigation of the phenylethanolamine N-methyltransferase gene as
 RT a candidate gene for hypertension.";
 RT Hypertension 26:595-601(1995).
 RN [3]
 RP SEQUENCE OF 16-285 FROM N.A.
 RC TISSUE=Adrenal gland;
 RX MEDLINE=90135920; PubMed=2575695;
 RA Weisberg E.P., Baruchin A., Stachowiak M.K., Stricker E.M.,
 RA Zigmund M.J., Kaplan B.B.;
 RT "Isolation of a rat adrenal cDNA clone encoding phenylethanolamine N-
 RT methyltransferase and cold-induced alterations in adrenal PNMT mRNA
 RT and protein.";
 RT Brain Res. Mol. Brain Res. 6:159-166(1989).
 RN [4]
 RP SEQUENCE OF 26-285 FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=89183620; PubMed=2928117;
 RA Mezey E.;
 RT "Cloning of the rat adrenal medullary phenylethanolamine-N-
 RT methyltransferase.";
 RT Nucleic Acids Res. 17:2125-2125(1989).
 CC -1- FUNCTION: CONVERTS NORADRENALINE TO ADRENALINE.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + phenylethanolamine -
 CC S-adenosyl-L-homocysteine + N-methylphenylethanolamine.
 CC -1- PATHWAY: Catecholamine biosynthesis, last step.
 CC -1- SIMILARITY: BELONGS TO THE PNMT/PNMT/TEMT FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X75333; CAA53082.1; -
 DR EMBL; U11694; AAA91779.1; -
 DR EMBL; U11275; AAA91779.1; JOINED.
 DR EMBL; X14211; CAA32428.1; -
 DR PIR; A60060; A60060.
 DR PIR; S03614; S03614.
 DR InterPro; IPR001601; Methyltransf.
 DR InterPro; IPR000940; PNMT_PNMT_TEMTfm.
 DR Pfam; PF01234; PNMT_PNMT_TEMT; 1.
 DR PROSITE; PS01100; PNMT_PNMT_TEMT; 1.
 KW Transferase; Methyltransferase; Catecholamine biosynthesis.
 FT CONFLICT 16 17 DS -> LA (IN REF. 3).
 FT CONFLICT 32 32 D -> H (IN REF. 4).
 FT CONFLICT 204 204 T -> A (IN REF. 4).
 FT CONFLICT 211 211 N -> H (IN REF. 2 AND 3).
 FT CONFLICT 214 214 F -> L (IN REF. 2 AND 3).
 FT CONFLICT 285 285 V -> A (IN REF. 2).
 SO SEQUENCE 285 AA; 31670 MW; 28A239A5411AD26 CRC64;
 Query Match 96.6%; Score 28; DB 1; Length 285;
 Best Local Similarity 83.3%; Pred. No. 35;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EIRDLR 6
 DB 250 EIRDLR 255
 RESULT 10
 GRAP_MOUSE STANDARD; PRT; 430 AA.
 ID GRAP_MOUSE
 AC P03995;
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE G1a1 fibrillary acidic protein, astrocyte (GRAP).
 GN GRAP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85297756; PubMed=2994002;
 RA Balcarek J.M., Cowan N.J.;
 RT "Structure of the mouse g1a1 fibrillary acidic protein gene:
 RT implications for the evolution of the intermediate filament multigene
 RT family.";
 RT Nucleic Acids Res. 13:5527-5543(1985).
 RN [2]
 RP SEQUENCE OF 28-430 FROM N.A.
 RX MEDLINE=84194082; PubMed=6585825;
 RA Lewis S.A., Balcarek J.M., Krek V., Shelanski M., Cowan N.J.;
 RT "Sequence of a cDNA clone encoding mouse g1a1 fibrillary acidic
 RT protein: structural conservation of intermediate filaments.";
 RT Proc. Natl. Acad. Sci. U.S.A. 81:2743-2746(1984).
 RN [3]
 RP SEQUENCE OF 28-430 FROM N.A.
 RX MEDLINE=86101618; PubMed=3865511;
 RA Cowan N.J., Lewis S.A., Balcarek J.M., Krek V., Shelanski M.L.;
 RT "Structural implications of a cDNA clone encoding mouse g1a1
 RT fibrillary acidic protein.";
 RT Ann. N.Y. Acad. Sci. 455:575-582(1985).
 RN [4]
 RP REVISIONS TO N-TERMINUS.
 RX MEDLINE=90294716; PubMed=2163003;
 RA Benner M., Lampel K., Nakatani Y., Mall J., Banner C., Mearrow K.,
 RA Ohnaka M., Lipsky R., Fresse E.;
 RT "Characterization of human CDNA and genomic clones for g1a1

RT fibrillary acidic protein.";
 RN Brain Res. Mol. Brain Res. 7:277-286(1990).
 RN [5]
 RP SEQUENCE OF 1-26 FROM N.A.
 RC TISSUE-Embryo:
 RA MEDLINE-95074269; PubMed-7983160;
 RA Ralston J.E., Lu X., Hutcheson A.M., Quinlan R.A.;
 RT Identification of two N-terminal non-alpha-helical domain motifs
 RT important in the assembly of glial fibrillary acidic protein.";
 RL J. Cell Sci. 107:1935-1948(1994).
 CC -1- FUNCTION: GFAP, A CLASS-III INTERMEDIATE FILAMENT, IS A CELL-
 CC SPECIFIC MARKER THAT, DURING THE DEVELOPMENT OF THE CENTRAL-
 CC NERVOUS SYSTEM, DISTINGUISHES ASTROCYTES FROM OTHER GLIAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: K01347; AAA37678.1;
 DR EMBL: M25937; AAA37679.1;
 DR EMBL: X02801; CA26571.1; ALT_INT.
 DR EMBL: X78141; CA55020.1;
 DR PIR: B60052; VEMSGF.
 DR MGI: 95697; GFAP.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; Filament; 1.
 DR PROSITE: PS00226; IF; 1.
 DR KW Intermediate filament; Coiled coil.
 FT DOMAIN 1 69 HEAD.
 FT TIGR: M1381; ROD.
 FT DOMAIN 375 430 TAIL.
 FT DOMAIN 70 101 COIL 1A.
 FT DOMAIN 102 112 COIL 1B.
 FT DOMAIN 113 211 COIL 1B.
 FT DOMAIN 212 227 LINKER 12.
 FT DOMAIN 228 249 COIL 2A.
 FT DOMAIN 250 253 COIL 2B.
 FT DOMAIN 254 374 COIL 2B.
 FT CONFLICT 347 347 H -> Q (IN REF. 2 AND 3).
 FT SEQUENCE 430 AA; 49917 MW; 446554A65B82DFC9 CRC64;
 SQ
 Query Match 96.6%; Score 28; DB 1; Length 430;
 Best Local Similarity 83.3%; Pred. No. 55;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EIRDLR 6
 ID I:|||||
 DB 204 EVRDLR 209
 RESULT 11
 CAB2_METJA STANDARD; PRT; 618 AA.
 AC 058776;
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cardamoyl-phosphate synthase large chain, C-terminal section (EC
 DE 6.3.5.5) (Cardamoyl-phosphate synthetase ammonia chain).
 GN CARB2 OR M1381.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2861 / ATCC 43067;
 RX MEDLINE-96337999; PubMed-8688067;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Claydon R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.E., Fuhrman J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurt M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O - 2 ADP +
 CC phosphate + L-glutamate + carbamoyl phosphate.
 CC -1- CORFACTOR: Binds three manganese ions (By similarity).
 CC -1- PATHWAY: Arginine biosynthesis.
 CC -1- PATHWAY: Pyrimidine biosynthesis; first step.
 CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain
 CC promotes the hydrolysis of glutamate to ammonia, which is used by
 CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE CARB FAMILY, C-TERMINAL SECTION.
 CC -1- CAUTION: Sequence of carb is split into two genes in M.jannaschii
 CC (M1378 and M1381).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U67578; AAB9391.1;
 DR HSSP: P00968; IBXR.
 DR TIGR: M1381;
 DR InterPro: IPR005483; CPase_L.
 DR InterPro: IPR005479; CPase_L_D2.
 DR InterPro: IPR005480; CPase_L_D3.
 DR InterPro: IPR005481; CPase_L_N.
 DR InterPro: IPR004362; MGS-like.
 DR Pfam: PF00289; CPase_L_chain; 1.
 DR Pfam: PF02786; CPase_L_D2; 2.
 DR Pfam: PF02787; CPase_L_D3; 1.
 DR Pfam: PF02142; MGS; 1.
 DR PRINTS: PRO0098; CPASE.
 DR PROSITE: PS00866; CPASE_1; 1.
 DR PROSITE: PS00867; CPASE_2; 1.
 DR KW Arginine biosynthesis; Pyrimidine biosynthesis; Lysase; ATP-binding;
 KW Manganese; Complete proteome.
 FT DOMAIN 81 477 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
 FT DOMAIN 478 618 ALLOSTERIC DOMAIN.
 FT METAL 358 358 MANGANESE 3 (BY SIMILARITY).
 FT METAL 370 370 MANGANESE 3 (BY SIMILARITY).
 FT SEQUENCE 618 AA; 68453 MW; 8C3D4D9C192A954E CRC64;
 SQ
 Query Match 96.6%; Score 28; DB 1; Length 618;
 Best Local Similarity 83.3%; Pred. No. 81;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EIRDLR 6
 ID I:|||||
 DB 33 EVRDLR 38
 RESULT 12
 LAMA_CHICK
 ID LAMA_CHICK STANDARD; PRT; 657 AA.
 AC P13648;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lamin A.
 OS Gallus gallus (Chicken).

OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
 OC Archaeopteryx: Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90012208; PubMed=2795656;
 RA Peter M., Kitten G.T., Lehner C.F., Vordbuerger K., Baller S.M.,
 RA Maridor G., Nigg E.A.;
 RT Cloning and sequencing of cDNA clones encoding chicken lamins A and
 RT B1 and comparison of the primary structures of vertebrate A- and
 RT B-type lamins.
 RL J. Mol. Biol. 208:393-404(1998).
 CC -1- FUNCTION: LAMINS ARE COMPONENTS OF THE NUCLEAR LAMINA, A FIBROUS
 CC LAYER ON THE NUCLEOLAR PLASMIC SIDE OF THE INNER NUCLEAR MEMBRANE,
 CC WHICH IS THOUGHT TO PROVIDE A FRAMEWORK FOR THE NUCLEAR ENVELOPE
 CC AND MAY ALSO INTERACT WITH CHROMATIN.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X16879; CAA34762.1; -
 CC PIR: S05517; S05517.
 CC InterPro: IPR001322; IF_tail.
 CC InterPro: IPR001230; IF_tail.
 CC Pfam: PF00038; Filament; 1.
 CC DR PROSITE: PS00226; IF; 1.
 CC DR PROSITE: PS00226; IF; 1.
 CC KM Intermediate filament; Coiled coil; Nuclear protein; Lipoprotein;
 CC Prenylation; Phosphorylation; Alternative splicing.
 FT DOMAIN 1 32 HEAD.
 FT 1 32
 FT DOMAIN 33 382 ROD.
 FT 33 382
 FT DOMAIN 383 657 TAIL.
 FT 383 657
 FT DOMAIN 33 69 COIL 1A.
 FT 70 69
 FT DOMAIN 80 217 LINKER 1.
 FT 217 217
 FT DOMAIN 218 241 COIL 1B.
 FT 242 241
 FT DOMAIN 242 382 LINKER 2.
 FT 382 382
 FT LIPID 654 654 FARNESYL (BY SIMILARITY).
 SQ SEQUENCE 657 AA; 73164 MW; FD0B31828A77AF08 CRC64;
 Query Match 96.68; Score 28; DB 1; Length 657;
 Best Local Similarity 83.38; Pred. No. 87;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RP SEQUENCE FROM N.A.
 RC STRAIN=RL
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson M.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1".
 RL Science 286:1571-1577(1999).
 CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
 CC phosphate + L-glutamate + carbamoyl phosphate.
 CC -1- COFACTOR: Binds three manganese ions (by similarity).
 CC -1- PATHWAY: Arginine biosynthesis.
 CC -1- SUBUNIT: Pyrimidine biosynthesis; first step.
 CC promotes the hydrolysis of glutamine to ammonia, which is used by
 CC the large (or ammonia) chain to synthesize carbamoyl phosphate (by
 CC similarity).
 CC -1- SIMILARITY: BELONGS TO THE CARB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE001924; AAF10248.1; -
 CC HSSP: P00968; 1A9X.
 CC TIGR: DR0668;
 CC InterPro: IPR005483; CPase_L.
 CC InterPro: IPR005479; CPase_L.D2.
 CC InterPro: IPR005480; CPase_L.D3.
 CC InterPro: IPR005481; CPase_L.N.
 CC Pfam: PF02789; CPase_L-chain; 2.
 CC Pfam: PF02786; CPase_L.D2; 2.
 CC Pfam: PF02787; CPase_L.D3; 1.
 CC DR PRINTS: PR00098; CPASE.
 CC DR PROSITE: PS00866; CPASE_1; FALSE_NEG.
 CC DR PROSITE: PS00867; CPASE_2; 2.
 CC DR Arginine biosynthesis; Pyrimidine biosynthesis; Lysase; Repeat;
 CC ATP-binding; Manganese; Complete proteome.
 CC KM Arginine biosynthesis; Pyrimidine biosynthesis; Lysase; Repeat;
 CC ATP-binding; Manganese; Complete proteome.
 CC FT DOMAIN 1 402 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
 FT 403 546 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
 FT DOMAIN 547 929 ALLOSTERIC DOMAIN.
 FT 930 1024
 FT REPEAT 547 1024
 FT NP_BIND 153 210
 FT NP_BIND 303 353
 FT METAL 285 285
 FT METAL 299 299
 FT METAL 301 301
 FT METAL 822 822
 FT METAL 834 834
 SQ SEQUENCE 1024 AA; 11646 MW; A7CAB7A99C6B9715 CRC64;
 Query Match 96.68; Score 28; DB 1; Length 1024;
 Best Local Similarity 83.38; Pred. No. 1; Indels 0; Gaps 0;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

ID VGNB_RCMV STANDARD: PRT: 1864 AA.
 AC P35930:
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Genome polyprotein B (Contains: Protease cofactor; Membrane binding
 protein; VPg; Protease (EC 3.4.22.-); RNA polymerase (EC 2.7.7.48)).
 OS Red clover mottle virus (RCMV).
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Comoviridae;
 CC Comovirus.
 CC NCBI_TaxID=12262;
 RX NCBI_TaxID=12262;
 RP SEQUENCE FROM N.A.
 RC STRAIN-S:
 RA MEDLINE-93019077; PubMed-1402822;
 RA Shanks M., Lomonosoff G.P.,
 RT "The nucleotide sequence of red clover mottle virus bottom component
 RNA."
 RL J. Gen. Virol. 73:2473-2477(1992).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC (RNA)(N).
 CC -1- P1M: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: TO OTHER COMOVIRUSES GENOME POLYPROTEIN B.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X64886; GANA6104.1; -
 DR PIR: J01657; J01657.
 DR MEROPS: C03.003; -
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3d.
 DR Pfam: PF00680; RNA_dep_RNA_pol. 1.
 DR Pfam: PF00910; RNA_helicase. 1.
 KM Polyprotein; Transmembrane; Hydrolyase; Protease; Transferase;
 KW RNA-directed RNA polymerase; ATP-binding.
 FT CHAIN 1 315
 FT CHAIN 316 915
 FT CHAIN 916 943
 FT CHAIN 944 1151
 FT CHAIN 1152 1864
 FT NP_BIND 483 490
 FT ACT_SITE 983 983
 FT ACT_SITE 1019 1019
 FT ACT_SITE 1109 1109
 SQ SEQUENCE 1864 AA; 210255 MW; 8DD9D6C17C2B154 CRC64;
 Query Match 96.6%; Score 28; DB 1; Length 1864;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EIRDLR 6
 DB 1726 EVRDLR 1731
 RESULT 15
 PROA_PSEAE STANDARD: PRT: 421 AA.
 ID PROA_PSEAE
 AC Q9HX20;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Gamma-glutamyl phosphate reductase (GPR) (EC 1.2.1.41) (Glutamate-5-
 semialdehyde dehydrogenase) (GSA dehydrogenase).
 DE dehydrogenase). (GSA dehydrogenase).
 GN PROA OR PA4007.
 OS Pseudomonas aeruginosa.

CC Bacteria; Proteobacteria; gamma subphylum; Pseudomonadaceae;
 CC Pseudomonas.
 CC NCBI_TaxID=287;
 CC [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan X.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964(2000).
 CC -1- FUNCTION: CATALYZES THE NADPH DEPENDENT REDUCTION OF L-GAMMA-
 CC GLUTAMYL 5-PHOSPHATE INTO L-GLUTAMATE 5-SEMIALDEHYDE AND
 CC PHOSPHATE. THE PRODUCT SPONTANEOUSLY UNDERGOES CYCLIZATION TO FORM
 CC 1-PYRROLINE-5-CARBOXYLATE.
 CC -1- CATALYTIC ACTIVITY: L-glutamate 5-semialdehyde + phosphate +
 CC NADP(+) - L-glutamate 5-phosphate + NADPH.
 CC -1- PATHWAY: Proline biosynthesis; second step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE GAMMA-GLUTAMYL PHOSPHATE REDUCTASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE004817; AAG07394.1; -
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000965; Gglut_PP_reduct.
 DR Pfam: PF00171; aldehyd. 1.
 DR TIGRPFAM: TIGR00407; proA. 1.
 DR PROSITE: PS01223; PROA. 1.
 KW oxidoreductase; Proline biosynthesis; NADP; Complete proteome.
 SQ SEQUENCE 421 AA; 45044 MW; 447C9171BC5F35FD CRC64;
 Query Match 93.1%; Score 27; DB 1; Length 421;
 Best Local Similarity 83.3%; Pred. No. 90;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EIRDLR 6
 DB 99 EIRDMR 104

Search completed: November 13, 2002, 13:16:54
 Job time : 5.97872 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OW protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 : Search time 19,2766 Seconds
(without alignments)
64.134 Million cell updates/sec

Title: US-09-856-086-6
Perfect score: 29
Sequence: 1 EIRDLR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp Unclassified:*
- 15: sp_Virus:*
- 16: sp_Bacteriap:*
- 17: sp_Archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	99	12 Q9E869	Q9E869 gallid herp
2	29	100.0	99	12 Q9E861	Q9E861 gallid herp
3	29	100.0	109	10 Q9X1Y0	Q9X1Y0 cryza saliv
4	29	100.0	242	16 Q9CIG8	Q9CIG8 pasteurilla
5	29	100.0	338	12 Q8Q0Y1	Q8Q0Y1 turkey herp
6	29	100.0	338	12 Q8Q0X7	Q8Q0X7 turkey herp
7	29	100.0	339	12 Q9DGW5	Q9DGW5 turkey herp
8	29	100.0	339	12 Q9J306	Q9J306 marek's dis
9	29	100.0	339	12 Q8Q0Y2	Q8Q0Y2 turkey herp
10	29	100.0	339	12 Q8Q0X9	Q8Q0X9 turkey herp
11	29	100.0	339	12 Q8Q0X8	Q8Q0X8 turkey herp
12	29	100.0	339	12 Q8Q0X6	Q8Q0X6 turkey herp
13	29	100.0	339	12 Q8Q0X5	Q8Q0X5 turkey herp
14	29	100.0	339	12 Q8Q0X4	Q8Q0X4 turkey herp
15	29	100.0	347	2 Q47465	Q47465 erwinia car
16	29	100.0	347	2 Q47465	Q47465 erwinia car

17	29	100.0	347	2 Q47471	Q47471 erwinia car
18	29	100.0	377	16 Q9K6U6	Q9K6U6 bacillus ha
19	29	100.0	386	4 Q8TCR7	Q8TCR7 homo saplen
20	29	100.0	394	12 Q9JGK4	Q9JGK4 turkey herp
21	29	100.0	398	12 Q67633	Q67633 marek's dis
22	29	100.0	399	17 Q8TT74	Q8TT74 methanocarc
23	29	100.0	408	17 Q8ZVU9	Q8ZVU9 pyrobaculum
24	29	100.0	412	5 Q9Y0H1	Q9Y0H1 tetrahymena
25	29	100.0	467	3 Q96VC3	Q96VC3 mucor circi
26	29	100.0	497	3 Q9P3C2	Q9P3C2 neurospora
27	29	100.0	938	16 Q9HIX7	Q9HIX7 pseudomonas
28	28	96.6	133	10 Q9XIK7	Q9XIK7 arabidopsis
29	28	96.6	220	17 Q26792	Q26792 methanobact
30	28	96.6	254	17 Q8TUR7	Q8TUR7 methanopyru
31	28	96.6	268	10 Q9SG74	Q9SG74 arabidopsis
32	28	96.6	338	10 Q9CAF5	Q9CAF5 arabidopsis
33	28	96.6	347	16 Q9KGN0	Q9KGN0 bacillus ha
34	28	96.6	618	17 Q58776	Q58776 caenorhabdi
35	28	96.6	633	5 Q01976	Q01976 methanococc
36	28	96.6	752	17 Q8TVU9	Q8TVU9 methanopyru
37	28	96.6	1024	16 Q9RMR0	Q9RMR0 delinococcus
38	28	96.6	1193	16 Q9PH89	Q9PH89 xyloella fas
39	28	96.6	1206	3 Q9C1F2	Q9C1F2 pichia angu
40	27	93.1	244	17 Q8TYB5	Q8TYB5 methanopyru
41	27	93.1	290	12 Q83865	Q83865 nllaparvala
42	27	93.1	304	16 Q8Y4T0	Q8Y4T0 listeria mo
43	27	93.1	309	16 Q9RUF6	Q9RUF6 delinococcus
44	27	93.1	322	16 Q8X152	Q8X152 clostridium
45	27	93.1	323	5 Q9W1F7	Q9W1F7 drosophila

ALIGNMENTS

RESULT 1

Q9E869 PRELIMINARY: PRT: 99 AA.

AC Q9E869: 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Meg (Fragment).

GN MEO.

OS Gallid herpesvirus 1.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.

OX NCBI_TaxID=10386;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20481657; PubMed=11024146;

RA Yamaguchi T., Kaplan S.L., Wakenell P., Schat K.A.;

RT "Transactivation of latent Marek's disease herpesvirus genes in QT35,

RT a quail fibroblast cell line, by herpesvirus of turkeys.";

RL J. Virol. 74:10176-10186(2000).

DR EMBL: AF193002; AAC27971.1; -

DR InterPro: IPR000837; Leuzip.fos.

DR InterPro: IPR004827; TF_bZIP.

DR Pfam: PF00170; bZIP.1.

DR PRINTS: PRO0042; LEUZIPRPOS.

DR SMART: SM00338; BRLZ.1.

FT NON_TER 1

FT NON_TER 99

SEQUENCE 99 AA: 11869 MW: 0399D3ADFB2B5A189 CRC64;

Query Match 100.0%; Score 29; DB 12; Length 99;

Best Local Similarity 100.0%; Pred. No. 72;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6

DB 70 EIRDLR 75

RESULT 2

09E861 PRELIMINARY; PRT; 99 AA.
ID 09E861
AC 09E861;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Meg protein (Fragment).
GN MEO.
OS Gallid herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
NCBI_TaxID=10386;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV1988;
RX MEDLINE=20481657; PubMed=11024146;
RA Yamanuchi T., Kaplan S.L., Wakenell P., Schat K.A.;
RT "Transactivation of latent Marek's disease herpesvirus genes in Q735,
a quail fibroblast cell line, by herpesvirus of turkeys.";
J. Virol. 74:10176-10186(2000).
DR EMBL; AF193011; AAC29915.1; -;
DR InterPro; IPR000837; Leuzip_Fos.
DR Pfam; PF00170; bZIP.1
DR PRINTS; PR00042; LEUZIPPFFOS.
DR SMART; SM00338; BRLZ; 1.
FT NON_TER 1
FT TER 99
SQ SEQUENCE 99 AA; 11837 MW; 1BF9D3B9667CD8AB CRC64;

Query Match

Best Local Similarity 100.0%; Score 29; DB 12; Length 99;
Pred. No. 72;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
| | | | |

DB 70 EIRDLR 75

RESULT 3

09X1Y0 PRELIMINARY; PRT; 109 AA.
ID 09X1Y0
AC 09X1Y0;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. NIPONBARE;
RA Sasaki T., Nagamura Y., Yamamoto K.;
RT "Oryza sativa niponbare(GA3) genomic DNA, chromosome 6, PAC
clone: p0681f10.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB026295; BAA81850.1; -;
KW Hypothetical protein.
SQ SEQUENCE 109 AA; 12242 MW; 1438B8CA8813ADF9 CRC64;

Query Match

Best Local Similarity 100.0%; Score 29; DB 10; Length 109;
Pred. No. 79;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
| | | | |

DB 66 EIRDLR 71

RESULT 4

09C1G8 PRELIMINARY; PRT; 242 AA.
ID 09C1G8
AC 09C1G8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein PM1266.
GN PM1266.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang O., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70.";
Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006165; AAK0350.1; -;
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transport.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transport; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 242 AA; 27098 MW; 17D821923C156E86 CRC64;

Query Match

Best Local Similarity 100.0%; Score 29; DB 16; Length 242;
Pred. No. 17e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
| | | | |

DB 5 EIRDLR 10

RESULT 5

08QOY1 PRELIMINARY; PRT; 338 AA.
ID 08QOY1
AC 08QOY1;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MEO.
GN MEO.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-814;
RA Wei P., Cui Z.;
RT "The comparison of meg gene sequences amplified from different
RT pathotypes of Marek's disease virus.";
RL Chin. J. Prev. Vet. Med. 24:88-92(2002).
DR EMBL; AF493551; AAL99997.1; -;
KW Hypothetical protein.
SQ SEQUENCE 338 AA; 37126 MW; FD039B3A25DB438A CRC64;

Query Match

Best Local Similarity 100.0%; Score 29; DB 12; Length 338;
Pred. No. 2.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EIRDLR 6
| | | | |

DB 102 EIRDLR 107

RESULT 6

08QOY7

ID	0800X7	PRELIMINARY;	PRF;	338 AA.
AC	0800X7			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	MEQ.			
GN	MEQ.			
OS	Turkey herpesvirus.			
OC	viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
CC	Alphaherpesvirinae; Marek's disease-like viruses.			
OX	NCBI_TaxID=10390;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-CV1988;			
RA	Wel P., Cul Z.;			
RT	"The comparison of meq gene sequences amplified from different			
RT	pathotypes of Marek's disease virus.";			
RL	Chin. J. Prev. Vet. Med. 24:88-92(2002).			
DR	EMBL: AF493555; AAM00002.1; -			
SQ	SEQUENCE 338 AA; 37154 MW; ABB10E1D8649872 CRC64;			
Query Match	100.0%;	Score 29;	DB 12;	Length 338;
Best Local Similarity	100.0%;	Pred. No. 2,4e+02;		
Matches 6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Db	1 EIRDLR 6			
Db	102 EIRDLR 107			
RESULT 7				
09DGM5	PRELIMINARY;	PRT;	339 AA.	
AC	09DGM5;			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	MEQ protein.			
GN	MDV005 OR MDV076.			
OS	Turkey herpesvirus.			
OC	viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
CC	Alphaherpesvirinae; Marek's disease-like viruses.			
OX	NCBI_TaxID=10390;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-MD5;			
RX	MEDLINE-20392152; PubMed-10933706;			
RA	Tulman E.R., Afonso C.L., Lu Z., Zsak L., Rock D.L., Kutish G.F.;			
RT	"The genome of a very virulent Marek's disease virus.";			
RL	J. Virol. 74:7980-7988(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-MD5;			
RA	Tulman E.R., Afonso C.L., Lu Z., Zsak L., Rock D.L., Kutish G.F.;			
CC	Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.			
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE BZIP FAMILY.			
DR	EMBL: AF243438; AAG14278.1; -			
DR	EMBL: AF243438; AAG14255.1; -			
DR	InterPro: IPR000837; leuzip_pos.			
DR	InterPro: IPR002965; P_tlc_extensions.			
DR	InterPro: IPR004827; TF_bzip.			
DR	Pfam: PF001170; bzip.1.			
DR	PRINTS: PRO0042; LEUZIPPROS.			
DR	PRINTS: PRO1317; PRICHEXTENSN.			
DR	SMART: SM00338; BRZ.1.			
DR	PROSITE: PS00036; BZIP_BASIC.1.			
KW	DNA-binding; Nuclear protein.			
SQ	SEQUENCE 339 AA; 37224 MW; D2B1E34450B8463A CRC64;			
Query Match	100.0%;	Score 29;	DB 12;	Length 339;
Best Local Similarity	100.0%;	Pred. No. 2,4e+02;		
Matches 6;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

OY	1	EIRDLR	6		
Db	102	EIRDLR	107		
RESULT 8					
069306					
ID	069306	PRELIMINARY:	PRT:	339	AA.
AC	069306:				
DT	01-NOV-1996	(TREMBLrel. 01, Created)			
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)			
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)			
DE	MDV_ECO_Q protein (MEQ).				
GN	R-LORF.				
OS	Marek's disease herpesvirus (strain GA) (MDHV), and				
OS	Turkey herpesvirus.				
OC	Viruses: dsDNA viruses, no RNA stage; Herpesviridae;				
OC	Alphaherpesvirinae; Marek's disease-like viruses.				
OX	NCBI_TaxID=10388, 10390:				
RN	[1]				
RE	SEQUENCE FROM N.A.				
RC	SPECIES=gaillid herpesvirus 1, and Turkey herpesvirus; STRAIN-GA;				
RA	MEDLINE=92237304; PubMed=1315048;				
RA	Jones D., Lee L., Liu J.L., Kung H.J., Tilletson J.K.;				
RT	"Marek disease virus encodes a basic-leucine zipper gene resembling				
RT	the fos/jun oncogenes that is highly expressed in lymphoblastoid				
RT	tumors";				
RL	Proc. Natl. Acad. Sci. U.S.A. 89:4042-4046(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=turkey herpesvirus; STRAIN-GA;				
RA	Lee L.F., Mu P., Sul D., Ren D., Kung H.J.;				
RL	"The Complete UL Sequence of Serotype I Marek's Disease Virus";				
RL	Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2000).				
CC	-1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).				
CC	-1- SIMILARITY: BELONGS TO THE BZIP FAMILY.				
DR	EMBL: M89471; AAB48631.1; "				
DR	EMBL: AF147806; AAF67210.1; "				
DR	EMBL: AF147806; AAF67298.1; "				
DR	InterPro: IPR004827; TF_bzip.				
DR	Pfam: PF00170; BZIP_1.				
DR	SMART: SM00338; BRL1_1.				
DR	PROSITE: PS00036; BZIP_BASIC; 1.				
KW	DNA-binding; Nuclear protein.				
FT	VARIANT 35	K -> R.			
FT	VARIANT 77	K -> E.			
SO	SEQUENCE 339 AA; 37234 MW; DF16B8F68C8463A CRC64;				
Query Match		100.0%; Score 29; DB 12; Length 339;			
Best Local Similarity		100.0%; Pred. No. 2,4e+02;			
Matches	6; Conservative	0; Mismatches	0; Indels	0; Gaps	0;
OY	1	EIRDLR	6		
Db	102	EIRDLR	107		
RESULT 9					
080072					
ID	080072	PRELIMINARY:	PRT:	339	AA.
AC	080072:				
DT	01-JUN-2002	(TREMBLrel. 21, Created)			
DT	01-JUN-2002	(TREMBLrel. 21, Last sequence update)			
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)			
DE	MEQ.				
GN	MEQ.				
OS	Turkey herpesvirus.				
OC	Viruses: dsDNA viruses, no RNA stage; Herpesviridae;				
OC	Alphaherpesvirinae; Marek's disease-like viruses.				
OX	NCBI_TaxID=10390:				
RN	[1]				
RP	SEQUENCE FROM N.A.				

RC STRAIN-0093;
RA Wei P., Cui Z.;
RT "The comparison of med gene sequences amplified from different
pathotypes of Marek's disease virus.";
RL Chin. J. Prev. Vet. Med. 24:88-92(2002).
DR EMBL: AF493550; AAL99996.1; -.
SQ SEQUENCE 339 AA; 37381 MW; F049E9EB1426ABCA CRC64;

Query Match 100.0%; Score 29; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
Db 102 EIRDLR 107

RESULT 10
ID 0800Y0 PRELIMINARY; PRT; 339 AA.
AC 0800Y0;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE MEQ.
GN MEQ.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-0095;
RA Wei P., Cui Z.;
RT "The comparison of med gene sequences amplified from different
pathotypes of Marek's disease virus.";
RL Chin. J. Prev. Vet. Med. 24:88-92(2002).
DR EMBL: AF493552; AAL99998.1; -.
SQ SEQUENCE 339 AA; 37302 MW; DC333E5172F9FB11 CRC64;

Query Match 100.0%; Score 29; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
Db 102 EIRDLR 107

RESULT 11
ID 0800X9 PRELIMINARY; PRT; 339 AA.
AC 0800X9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE MEQ.
GN MEQ.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-0297;
RA Wei P., Cui Z.;
RT "The comparison of med gene sequences amplified from different
pathotypes of Marek's disease virus.";
RL Chin. J. Prev. Vet. Med. 24:88-92(2002).
DR EMBL: AF493553; AAL99999.1; -.
SQ SEQUENCE 339 AA; 37156 MW; AA71C0CF3E7C183A CRC64;

Query Match 100.0%; Score 29; DB 12; Length 339;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
Db 102 EIRDLR 107

RESULT 12
ID 0800X8 PRELIMINARY; PRT; 339 AA.
AC 0800X8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE MEQ.
GN MEQ.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-0304;
RA Wei P., Cui Z.;
RT "The comparison of med gene sequences amplified from different
pathotypes of Marek's disease virus.";
RL Chin. J. Prev. Vet. Med. 24:88-92(2002).
DR EMBL: AF493554; AAM00001.1; -.
SQ SEQUENCE 339 AA; 37494 MW; 9B56B0AC7FF9CD5 CRC64;

Query Match 100.0%; Score 29; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
Db 102 EIRDLR 107

RESULT 13
ID 0800X6 PRELIMINARY; PRT; 339 AA.
AC 0800X6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE MEQ.
GN MEQ.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-G2;
RA Wei P., Cui Z.;
RT "The comparison of med gene sequences amplified from different
pathotypes of Marek's disease virus.";
RL Chin. J. Prev. Vet. Med. 24:88-92(2002).
DR EMBL: AF493556; AAM00003.1; -.
SQ SEQUENCE 339 AA; 37229 MW; 1C40566CC45FBCA CRC64;

Query Match 100.0%; Score 29; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EIRDLR 6
Db 102 EIRDLR 107

RESULT 14

```

0800X5
ID 0800X5 PRELIMINARY; PRT; 339 AA.
AC 0800X5;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE MEQ.
GN MEQ.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_Taxid=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N;
RA Wel P., Cui Z.;
RT "The comparison of meq gene sequences amplified from different
RL pathotypes of Marek's disease virus."
RL Chin. J. Prev. Vet. Med. 24:88-92(2002).
DR EMBL: AF493557; AAM0004.1; -.
SQ SEQUENCE 339 AA; 3728 MW; 4642F4EB12F9E7FF CRC64;

```

```

Query Match 100.0%; Score 29; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EIRDLR 6
DB 102 EIRDLR 107

```

RESULT 15

```

0800X4
ID 0800X4 PRELIMINARY; PRT; 339 AA.
AC 0800X4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE MEQ.
GN MEQ.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Marek's disease-like viruses.
OX NCBI_Taxid=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=648A;
RA Wel P., Cui Z.;
RT "The comparison of meq gene sequences amplified from different
RL pathotypes of Marek's disease virus."
RL Chin. J. Prev. Vet. Med. 24:88-92(2002).
DR EMBL: AF493558; AAM0005.1; -.
SQ SEQUENCE 339 AA; 3726 MW; CEAB6F3B79474527 CRC64;

```

```

Query Match 100.0%; Score 29; DB 12; Length 339;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 EIRDLR 6
DB 102 EIRDLR 107

```

Search completed: November 13, 2002, 13:16:11
 Job time : 21.2766 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 39.6596 Seconds
(Without alignments)
26.879 Million cell updates/sec

Title: US-09-856-086-7

Perfect score: 39
Sequence: 1 EGEERDLR 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq_101002:*

- 1: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
- 2: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
- 3: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
- 4: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
- 5: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
- 6: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
- 7: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
- 8: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
- 9: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
- 10: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
- 11: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
- 12: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
- 13: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
- 14: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
- 15: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
- 16: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
- 17: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
- 18: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
- 19: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
- 20: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
- 21: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
- 22: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
- 23: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	8	21	Test antigen #8 fo
2	39	100.0	441	19	Human neurofilamen
3	39	100.0	543	22	Human polypeptide,
4	36	92.3	640	21	Human mitochondria
5	35	89.7	403	23	Mouse ischaemic co
6	34	87.2	1093	22	ABR08201
7	34	87.2	1975	22	ABR62094
8	33	84.6	14	22	Human lipid metabo
9	33	84.6	85	20	Human peptide #201
10	33	84.6	85	20	HIV-1 group O iso1
					AAV05631
					AAV05637

11	33	84.6	455	22	ABR61289
12	33	84.6	1212	22	ABR71114
13	32	82.1	81	19	ABP33718
14	32	82.1	107	19	AAV86013
15	32	82.1	183	22	AAU37575
16	32	82.1	301	22	AAV64043
17	32	82.1	475	22	AAV93229
18	32	82.1	477	22	AAV79279
19	32	82.1	557	19	AAV20666
20	32	82.1	628	22	AAO21847
21	32	82.1	749	22	ABR65445
22	32	82.1	789	22	ABR65404
23	32	79.5	35	22	ABR9483
24	31	79.5	394	22	ABG26375
25	31	79.5	482	22	ABG08132
26	31	79.5	625	22	ABG01592
27	31	79.5	631	22	AAU15923
28	31	79.5	716	22	ABR66983
29	31	79.5	811	22	ABR63560
30	31	79.5	987	22	ABG01594
31	31	79.5	1177	22	AAV96721
32	31	79.5	1303	22	ABG12230
33	30	76.9	142	22	AAU03636
34	30	76.9	143	21	AAV03892
35	30	76.9	148	23	ABP00247
36	30	76.9	159	23	ABR77489
37	30	76.9	169	21	AAV07988
38	30	76.9	180	21	AAV04331
39	30	76.9	182	21	AAV07987
40	30	76.9	190	22	AAV90883
41	30	76.9	206	22	ABR60636
42	30	76.9	234	22	AAV66926
43	30	76.9	278	21	AAV04330
44	30	76.9	282	23	ABR91609
45	30	76.9	286	21	AAV04329

ALIGNMENTS

RESULT 1
AAB05932 standard; peptide: 8 AA.
ID AAB05932; AAB05932;
AC 17-OCT-2000 (first entry)
XX Test antigen #8 for spongiform and demyelinating disease diagnosis.
XX DE Human; cow; myelin; myelin neurofilament; immunogen; antigen;
XX KW bovine spongiform encephalopathy; BSE; multiple sclerosis;
XX KM Creutzfeldt-Jacob disease; CJD; demyelinating disease; diagnostic test.
XX OS Bos taurus.
XX OS Homo sapiens.
XX PN WO200031545-A1.
XX PD 02-JUN-2000.
XX PF 25-NOV-1999; 99WO-GB03936.
XX PR 26-NOV-1998; 98GB-0025948.
XX (UNLO) KING'S COLLEGE.
XX Ebringer A;
XX WPI; 2000-400194/34.
XX Diagnosing spongiform or demyelinating disease in vertebrates such as
XX bovine spongiform encephalopathy and Creutzfeldt-Jacob disease comprises

PT assaying a biological sample for myelin and/or myelin neurofilament
 PT antibodies -
 XX
 PS Claim 5; Page 2; 16pp; English.
 XX
 CC The present peptide may be used as a test antigen in a kit for diagnosing
 CC sporngiform or demyelinating disease in vertebrates, including bovine
 CC sporngiform encephalopathy (BSE), multiple sclerosis (MS) and antigenic
 CC Creutzfeldt-Jacob disease (CJD). Peptides used in the kit are antigenic
 CC components of myelin or myelin neurofilaments. Biological samples
 CC are assayed for antibodies, especially Iga antibodies, which bind to
 CC myelin and/or myelin neurofilaments or an antigenic peptide, such as the
 CC present sequence. Any reading in excess of two standard deviations of
 CC the healthy controls would indicate a positive response.
 CC
 SO Sequence 8 AA;
 Query Match 100.0%; Score 39; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EOEIRDLR 8
 DB 1 EOEIRDLR 8
 XX
 XX
 DE Human neurofilament-L wild type protein fragment 2.
 XX
 KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GRAP; p53; semaphorin III; HMPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMG-C; NSP-A;
 KW high mobility group protein C; neuroendocrine specific protein A.
 KW
 XX
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 DR WPI; 1998-609901/51.
 DR N-PSDB; AAX75758.
 XX
 XX Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 7; 258pp; English.

XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HMPF-1, high mobility group
 CC protein-C (HMG-C) and neuroendocrine specific protein A.
 CC
 SO Sequence 441 AA;
 Query Match 100.0%; Score 39; DB 19; Length 441;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EOEIRDLR 8
 DB 177 EOEIRDLR 184
 XX
 XX
 DE Human polypeptide, SEQ ID NO: 3132.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX
 OS Homo sapiens.
 XX
 PN EP130094-A2.
 XX
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 DR N-PSDB; AAK94387.
 XX
 XX 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 PT
 PS Claim 8; SEQ ID NO 3132; 1380pp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by

CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

SO Sequence 543 AA;

Query Match 100.0%; Score 39; DB 22; Length 543;
 Best Local Similarity 100.0%; Pred. No. 28;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EOEIRDLR 8
 |||||

Db 140 EOEIRDLR 147

RESULT 4

AY80296
 ID AY80296 standard; Protein: 640 AA.

AC AY80296;

DT 31-MAY-2000 (first entry)

DE Human mitochondrial phosphoenolpyruvate carboxylkinase SEQ ID NO:1.

XX Human; mitochondrial phosphoenolpyruvate carboxylkinase; PEPC-K; PK2;
 KM PEPC-K-mitochondrial; mPEPC; antisense oligonucleotide; modulation;
 KW phosphochloster; inhibition; diagnosis.

XX Homo sapiens.

OS US6030837-A.

PN 29-FEB-2000.

PD 03-AUG-1999; 99US-0366257.

PF 03-AUG-1999; 99US-0366257.

PR 03-AUG-1999; 99US-0366257.

PA (ISIS-) ISIS PHARM INC.

PI Mckay R, Cowsett LM, Butler MM;

DR WPI: 2000-205209/18.

DR N-PSDB; AA295313.

PT New antisense compound targeted to a nucleic acid molecule encoding
 PT human mitochondrial phosphoenolpyruvate carboxylkinase useful for
 PT treating a human with a mitochondrial phosphoenolpyruvate

PT carboxylkinase-associated disease -

XX Example 13; Column 41-46; 32pp; English.

CC AA295330 to AA295359 represent antisense oligonucleotides targeted to a
 CC nucleic acid molecule encoding human mitochondrial phosphoenolpyruvate
 CC carboxylkinase (also known as PEPC-mitochondrial; PEPC-M; PK2 and
 CC mPEPC), where the oligonucleotide specifically hybridise with and
 CC inhibit the expression of human mPEPC. The antisense oligonucleotides
 CC can be used for inhibiting the expression of mPEPC in human cells or
 CC tissues in vitro and can also be used for treating an animal.
 CC particularly a human suspected of having or being prone to a condition
 CC or disease associated with expression of mPEPC. They can also be used
 CC in diagnostics and as research reagents in sandwich and other assays.
 CC The present sequence represents human mPEPC.

SO Sequence 640 AA;

Query Match

Best Local Similarity 92.3%; Score 36; DB 21; Length 640;
 75.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EOEIRDLR 8
 |||||

Db 604 EOEVRDLR 611

RESULT 5

ABB57243
 ID ABB57243 standard; Protein: 403 AA.

AC ABB57243;

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related protein sequence SEQ ID NO:650.

XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KM vasospastic ischaemia; ischaemic condition; ischaemic disease.

XX Mus musculus.

PN WO20018188-A2.

PD 22-NOV-2001.

PF 18-MAY-2001; 2001WO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

PA (UNIT-) UNIT NIHON SCHOOL JURIDICAL PERSON.

PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;

DR WPI: 2002-034733/04.

DR N-PSDB; AB199627.

PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
 PT expression levels of particular genes defined in the specification or
 PT by determining the expression profile of a gene group comprising these

PT genes -

PS Claim 2; Page 1635-1637; 2690pp; English.

CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (AB199202 to AB199912, encoding
 CC the protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The
 CC expression levels or expression profiles produced by these genes are
 CC used as an indicator when screening for ischaemic condition-improving
 CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914
 CC represent PCR primers for a mouse ischaemic condition related sequence,
 CC which are used in the exemplification of the present invention.

SO Sequence 403 AA;

Query Match

Best Local Similarity 89.7%; Score 35; DB 23; Length 403;
 75.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EOEIRDLR 8
 |||||

Db 175 EEEVRDLR 182

RESULT 6

ABB08201
 ID ABB08201 standard; Protein: 1093 AA.

XX ABB08201;
 AC 04-MAR-2002 (first entry)
 DT
 XX
 DE Human lipid metabolism enzyme-1 (LME-1).
 KW Human; LME-1; lipid metabolism enzyme-1; cytostatic; neuroprotective;
 KM immunosuppressive; anti-inflammatory; cardiovascular; gene therapy;
 KW enzyme therapy; cancer; neurological disorder; autoimmune disorder;
 KM inflammatory disorder; cardiovascular disorder.
 OS Homo sapiens.
 XX
 PN MO200185956-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 11-MAY-2001; 2001WO-US15210.
 XX
 PR 11-MAY-2000; 2000US-203511P.
 PR 25-MAY-2000; 2000US-207903P.
 PR 07-JUN-2000; 2000US-210150P.
 PR 23-JUN-2000; 2000US-213392P.
 XX
 PA (INCYTE GENOMICS INC.
 XX
 PI Das D, Reddy R, Yao MG, Nguyen DB, Lu Y, Tribouley CM, Yue H;
 PI Khan FA, Gandhi AR, Au-young J, Lal P, Kearney L, Elliott VS;
 PI Ding L, Thornton M;
 DR N-PSDB; ABA96001.
 DR
 XX
 PT New lipid metabolism enzymes useful for diagnosing, treating and
 PT preventing cancer, neurological disorders, autoimmune/inflammatory
 PT disorders, and cardiovascular disorders
 XX
 PS Claim 1: Page 106-108; 122pp; English.
 XX
 CC The sequence represents a novel human lipid metabolism enzyme (LME-1) of
 CC the invention. The invention relates to novel human LME's, and the
 CC polynucleotides which identify and encode them. The enzymes of the
 CC invention have cytostatic, neuroprotective, immunosuppressive,
 CC anti-inflammatory, and cardiovascular activity. The polypeptides and
 CC polynucleotides have a use in gene therapy and enzyme therapy. The lipid
 CC metabolism enzymes are useful in the diagnosis, treatment and prevention
 CC of cancer, neurological disorders, autoimmune/inflammatory disorders, and
 CC cardiovascular disorders, and in the assessment of the effects of
 CC exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of lipid metabolism enzymes. LMEs may also be used to screen
 CC for compounds that modulate the activity of LME. Polynucleotides encoding
 CC LME may be used for somatic or germline gene therapy, to detect and
 CC quantify gene expression in biopsied tissues in which expression of LME
 CC may be correlated with disease, to generate a transcript image of a
 CC tissue or cell type, to generate hybridization probes useful in mapping
 CC the naturally occurring genomic sequence, and for screening libraries of
 CC compounds in drug screening techniques. The polypeptide sequences may be
 CC used to analyse the proteome of a tissue or cell type. Oligonucleotide
 CC primers derived from polynucleotide sequences encoding LME may be used to
 CC detect single nucleotide polymorphisms.
 CC
 SQ Sequence 1093 AA:
 XX
 Query Match 87.2%; Score 34; DB 23; Length 1093;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EOEIRDL 7
 |||||
 DB 934 EOEIRDL 940

RESULT 7
 ABB62094
 ID ABB62094 standard; Protein; 1975 AA.
 XX
 AC ABB62094;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 13074.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 OS Drosophila melanogaster.
 XX
 PN MO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR N-PSDB; ABL06197.
 DR
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions
 XX
 PS Disclosure; SEQ ID NO 13074; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 1975 AA:
 XX
 Query Match 87.2%; Score 34; DB 22; Length 1975;
 Best Local Similarity 87.5%; Pred. No. 8.7e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EOEIRDL 8
 |||||
 DB 1842 EOEIRDL 1849
 RESULT 8
 AAM98744
 ID AAM98744 standard; Peptide; 14 AA.
 XX
 AC AAM98744;
 XX
 DT 24-JAN-2002 (first entry)
 DE Human peptide #2019 encoded by a SNP oligonucleotide.
 KW Immunosuppressive; immunostimulatory; anti-inflammatory; cytostatic;
 KM neuroprotective; antimicrobial; gene therapy; vaccine; amyloid; cancer;
 KW amyloid protein; angiotensin; apoptosis related protein; cadherin;

XX WPI: 1996-132255/1.
 DR N-PSDB; AAX25181.
 XX New isolated HIV-1 group O strains - used to produce
 PT polynucleotides, antigens and antibodies for use in diagnosis and in
 PT vaccines for prevention of HIV-1 infection
 XX
 PS Disclosure: Fig 8B; 162pp; English.
 CC The present sequence represents the Ypu protein of HIV-1 group O
 CC (Oullier) virus isolate MP311 (FABA), as deduced from part of the
 CC MP311 genome (see AAX25181). The invention relates to new HIV-1 group
 CC O antigens, especially envelope protein antigens (see AAY05346-625),
 CC and the use of these antigens, or nucleic acids encoding them (see
 CC AAX25154-80), in the diagnosis and prophylaxis of AIDS. They can be
 CC used as reagents for detecting HIV-1 group O infection and for
 CC differentiating different types of HIV-1 group O infection.
 CC Vaccines that provide protective immunity against HIV-1 infection,
 CC in particular against HIV-1 group O infection, comprise at least
 CC one HIV-1 type O antigen, a nucleic acid encoding such an antigen,
 CC a virus-like particle comprising such an antigen, or an attenuated
 CC form of an HIV-1 type O strain. The invention also relates to new
 CC HIV-1 group O strains, obtained from patients from Cameroon, Gabon,
 CC Tchad, Nigeria, Senegal (such as MP311) and Niger.

SO Sequence 85 AA;

Query Match 84.6%; Score 33; DB 20; Length 85;
 Best Local Similarity 85.7%; Pred. No. 57;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDL 7
 ||:||||

DB 66 EOEVRDL 72

RESULT 11

ABB61289 ID ABB61289 standard; Protein; 455 AA.

XX ABB61289;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 10659.

KW Drosophila; developmental biology; cell signalling; insecticide;
 pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PE 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PT 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, LI PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL05392.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure; SEQ ID NO 10659; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded DNA
 CC sequences (AB57737-AB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 455 AA;

Query Match 84.6%; Score 33; DB 22; Length 455;
 Best Local Similarity 75.0%; Pred. No. 3; 1e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDL 8
 ||:||||

DB 97 EOEIRDLX 104

RESULT 12

ABB71114 ID ABB71114 standard; Protein; 1212 AA.

XX ABB71114;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 40134.

KW Drosophila; developmental biology; cell signalling; insecticide;
 pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PE 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PT 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PI Venter JC, Adams M, LI PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL15217.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX Disclosure; SEQ ID NO 40134; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
 CC sequences (AB57737-AB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

Sequence 1212 AA;

Query Match 84.6%; Score 33; DB 22; Length 1212;
 Best Local Similarity 87.5%; Pred. No. 8.2e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 EOEIRDLR 8
 |||||
 DB 625 EOEISDLR 632

RESULT 13
 ABP33718
 ID ABP33718 standard; Protein: 81 AA.

XX AC ABP33718;
 XX DF 08-JUL-2002 (first entry)
 XX DE Human transcription factor-like ORF2691 protein, SEQ ID NO:5382.

XX KM Human; ORF: open reading frame; ORF: drug screening; diagnosis;
 KM disease monitoring; cytokine; cell proliferation; cell differentiation;
 KM immune modulation; haematopoiesis regulation; tissue growth;
 KM angiogenesis; activin; inhibin; chemotactic; chemokinetic; haemostatic;
 KM thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KM behaviour; cancer; proliferative disorder; neurological disorder;
 KM cardiovascular disease; immune system disorder; organ transplantation;
 KM tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KM hypothyroidism; cholesterol ester storage disease; infection; vulnerability;
 KM vasotropic; antipruritic; antidiabetic; cytosolic; neurotrophic;
 KM neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KM cardiac; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
 KM dermatological; analgesic; virucide; antibacterial; fungicide.

XX OS Homo sapiens.
 XX PN MO200190366-AA2.
 XX PD 29-NOV-2001.
 XX PF 24-MAY-2001; 2001WO-US17076.
 XX PR 24-MAY-2000; 2000US-206690P.

XX PA (CURA-) CURAGEN CORP.
 XX PI Leach MD, Shinkets RA;
 XX N-PSDB; ABN77744.
 XX DR WPI: 2002-106200/14.
 XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation
 XX PS Claim 10; Page 1597; 2508pp; English.

XX CC sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80s identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85s identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/

chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 CC and antiinfective activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.

Sequence 81 AA;

Query Match 82.1%; Score 32; DB 23; Length 81;
 Best Local Similarity 85.7%; Pred. No. 83;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 EOEIRDL 7
 |||||
 DB 67 EOEIRDI 73

RESULT 14
 AAY66013
 ID AAY66013 standard; Protein: 107 AA.

XX AC AAY66013;
 XX DF 10-APR-2000 (first entry)
 XX DE S. pneumoniae derived protein #222.

XX KM Treatment; prevention; disease; diagnosis; gene therapy; screening;
 KM bacterial; antimicrobial; antibiotic; pathogenesis; infection.
 XX OS Streptococcus pneumoniae.
 XX PN WO9806734-A1.
 XX PD 19-FEB-1998.
 XX PF 15-AUG-1997; 97WO-US14436.
 XX PR 16-AUG-1996; 96US-0024022.
 XX PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX PI Black MT, Hodgson JE, Knowles DVC, Lonetto MA, Nicholas RO;
 XX Stodola RK;
 XX N-PSDB; AA296350.
 XX DR WPI: 1998-159452/14.
 XX PT Streptococcus pneumoniae proteins and related DNA - useful for
 PT screening compounds for antibacterial activity
 XX PS Claim 5; Page 501-502; 640pp; English.

XX CC This invention describes novel isolated Streptococcus pneumoniae
 CC polynucleotides (see AA296173-296494) and their encoded proteins (see
 CC AA296792-296182). The DNA, vectors and host cells described in the
 CC method of the invention are useful for the recombinant expression of the
 CC polypeptides. The polypeptides are useful for treatment or prevention of

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 11.5745 Seconds
(without alignments)
20.336 Million cell updates/sec

Title: US-09-856-086-7

Perfect score: 39

Sequence: 1 EQRDLR 8

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	74.4	29	1	US-08-475-989-5
2	29	74.4	29	2	US-08-475-985-5
3	29	74.4	29	2	US-08-472-172-17
4	29	74.4	29	3	US-08-256-839-5
5	29	74.4	35	4	US-08-750-624-12
6	29	74.4	37	4	US-09-106-568E-160
7	29	74.4	39	1	US-08-475-989-15
8	29	74.4	39	2	US-08-256-839-15
9	29	74.4	39	3	US-08-362-336A-15
10	29	74.4	65	4	US-09-111-470-8
11	29	74.4	238	4	US-07-803-622E-7
12	29	74.4	375	1	US-07-803-622E-9
13	29	74.4	375	1	US-07-803-622E-9
14	29	74.4	396	1	US-08-985-908-24
15	29	74.4	406	1	US-08-650-275-2
16	29	74.4	406	1	US-09-181-318-2
17	29	74.4	446	4	US-09-457-046B-74
18	29	74.4	466	4	US-09-610-401-3
19	29	74.4	466	4	US-09-610-401-4
20	29	74.4	693	2	US-08-380-403A-2
21	29	74.4	693	2	US-08-380-403A-5
22	29	74.4	693	2	US-08-895-628-2
23	29	74.4	693	2	US-08-895-628-5
24	29	74.4	693	4	US-08-895-810D-2
25	29	74.4	693	4	US-08-895-810D-5
26	29	74.4	725	3	US-09-234-337-12
27	29	74.4	739	1	US-07-803-622E-2

28	29	74.4	790	4	US-09-634-957-2	Sequence 2, Appl1
29	29	74.4	790	4	US-09-723-153-2	Sequence 2, Appl1
30	29	74.4	790	4	US-09-723-429-2	Sequence 2, Appl1
31	29	74.4	822	4	US-08-684-932A-58	Sequence 38, Appl1
32	29	74.4	2616	6	5206163-3	Patent No. 5206163
33	28	71.8	67	4	US-09-106-568E-152	Sequence 152, App
34	28	71.8	146	4	US-08-936-165A-349	Sequence 349, App
35	28	71.8	185	4	US-09-562-737-112	Sequence 112, App
36	28	71.8	317	2	US-08-726-306A-168	Sequence 168, App
37	28	71.8	348	1	US-08-843-521-6	Sequence 6, Appl1
38	28	71.8	348	2	US-08-953-041-2	Sequence 2, Appl1
39	28	71.8	348	4	US-09-159-417-2	Sequence 2, Appl1
40	28	71.8	348	4	US-09-012-871-6	Sequence 6, Appl1
41	28	71.8	370	4	US-09-134-001C-3769	Sequence 3769, Ap
42	28	71.8	394	4	US-09-218-197-2	Sequence 2, Appl1
43	28	71.8	410	1	US-08-471-033-40	Sequence 40, Appl1
44	28	71.8	410	1	US-08-471-033-43	Sequence 43, Appl1
45	28	71.8	410	2	US-08-471-044-40	Sequence 40, Appl1

ALIGNMENTS

RESULT 1
US-08-475-989-5
Sequence 5, Application US/08475989
Patent No. 5679352
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: RANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08475, 989
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256, 839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9202219, 3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-505 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-475-985-5

Query Match 74.4%; Score 29; DB 1; Length 29;
Best Local Similarity 62.5%; Pred. No. 27;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
:|||||:
DB 5 QOEFRDLK 12

RESULT 2
US-08-475-985-5
; Sequence 5, Application US/08475985
; Patent No. 5972349
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: SIA, Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: Synthetic Haemophilus Influenzae
; TITLE OF INVENTION: Conjugate Vaccine
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,985
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,839
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA93/00041
; FILING DATE: 03-FEB-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9202219.3
; FILING DATE: 03-FEB-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-506 MIS:Vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-475-985-5

Query Match 74.4%; Score 29; DB 2; Length 29;
Best Local Similarity 62.5%; Pred. No. 27;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
:|||||:
DB 5 QOEFRDLK 12

RESULT 3
US-08-472-172-17
; Sequence 17, Application US/08472172
; Patent No. 5985288
; GENERAL INFORMATION:
; APPLICANT: Munson, Jr., Robert S
; APPLICANT: Grass, Susan
; APPLICANT: Chong, Pele Y Y
; APPLICANT: Fahim, Raafat
; APPLICANT: SIA, Charles D Y
; APPLICANT: McVeery, Patrick
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Outer Membrane Protein P1 and Peptides
; TITLE OF INVENTION: Of Haemophilus Influenzae Type B
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,172
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/849,411
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24973
; REFERENCE/DOCKET NUMBER: 1038-471
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 063-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-472-172-17

Query Match 74.4%; Score 29; DB 2; Length 29;
Best Local Similarity 62.5%; Pred. No. 27;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
:|||||:
DB 5 QOEFRDLK 12

RESULT 4
US-08-256-839-5
; Sequence 5, Application US/08256839
; Patent No. 6018019
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: KANDIL, Ali
; APPLICANT: SIA, Charles

APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,839
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-373 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-256-839-5

Query Match 74.4%; Score 29; DB 3; Length 29;
Best Local Similarity 62.5%; Pred. No. 27;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 5 OOEFRDLK 12

RESULT 5
US-08-750-624-12
Sequence 12, Application US/08750624
Patent No. 6290971
GENERAL INFORMATION:
APPLICANT: Kandil, Ali
APPLICANT: James, Olive A
APPLICANT: Chong, Pele
APPLICANT: Klein, Michel H
TITLE OF INVENTION: ADJUVANT COMPOSITIONS COMPRISING A
TITLE OF INVENTION: MINERAL SALT AND ANOTHER IMMUNOSTIMULATING COMPOUND
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,624
FILING DATE: 26-FEB-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-660
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 35 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-750-624-12

Query Match 74.4%; Score 29; DB 4; Length 35;
Best Local Similarity 62.5%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 10 OOEFRDLK 17

RESULT 6
US-09-106-568E-160
Sequence 160, Application US/09106568E
Patent No. 6455248
GENERAL INFORMATION:
APPLICANT: Bhattacharjee, J.
APPLICANT: Suvaina, Kalavati
APPLICANT: Bhattacharjee, Vasker
TITLE OF INVENTION: METHODS AND REAGENTS FOR DETECTING FUNGAL PATHOGENS IN
TITLE OF INVENTION: A BIOLOGICAL SAMPLE
FILE REFERENCE: 96,247-A
CURRENT APPLICATION NUMBER: US/09/106,568E
CURRENT FILING DATE: 1998-06-29
PRIOR APPLICATION NUMBER: 08/650,809
PRIOR FILING DATE: 1997-05-20
NUMBER OF SEQ ID NOS: 160
SOFTWARE: Microsoft Word 97
SEQ ID NO 160
LENGTH: 37
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Polypeptide segment of Lys2-SCER shown in Figure 4.
US-09-106-568E-160

Query Match 74.4%; Score 29; DB 4; Length 37;
Best Local Similarity 71.4%; Pred. No. 35;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDL 7
DB 5 EREVRDL 11

RESULT 7
US-08-475-989-15
Sequence 15, Application US/08475989
Patent No. 5679352
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto

STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,989
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9202219.3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-505 MIS:V9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-475-989-15

Query Match 74.4%; Score 29; DB 1; Length 39;
Best Local Similarity 62.5%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 15 QOEFRDLK 22

RESULT 8
US-08-475-985-15
Sequence 15, Application US/08475985
Patent No. 5972349
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/475,985
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,839
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA93/00041
FILING DATE: 03-FEB-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9202219.3
FILING DATE: 03-FEB-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-506 MIS:V9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-475-985-15

Query Match 74.4%; Score 29; DB 2; Length 39;
Best Local Similarity 62.5%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 15 QOEFRDLK 22

RESULT 9
US-08-256-839-15
Sequence 15, Application US/08256839
Patent No. 6018019
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: KANDIL, Ali
APPLICANT: SIA, Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: Synthetic Haemophilus Influenzae
TITLE OF INVENTION: Conjugate Vaccine
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,839
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-373 MIS:Jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-256-839-15

Query Match
Best Local Similarity 74.4%; Score 29; DB 3; Length 39;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 15 QOEPRDLK 22

RESULT 10
US-09-362-336A-15
; Sequence 15, Application US/09362336A
; Patent No. 6440696
; GENERAL INFORMATION:
; APPLICANT: Band, Vimala
; APPLICANT: Gao, Qingshen
; TITLE OF INVENTION: No. 6440696e1 E6 Targeted Protein (E6TP1)
; FILE REFERENCE: 18475-017, E6TP1
; CURRENT APPLICATION NUMBER: US/09/362,336A
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: residues 963-1027 of GenBank Accession no. Z389009
US-09-362-336A-15

Query Match
Best Local Similarity 74.4%; Score 29; DB 4; Length 65;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 41 EEEVSLR 48

RESULT 11
US-09-111-470-8
; Sequence 8, Application US/09111470
; Patent No. 6277959
; GENERAL INFORMATION:
; APPLICANT: Valladeau, Jenny
; APPLICANT: Ravel, Odile
; APPLICANT: Bates, Elizabeth E.M.
; APPLICANT: Ford, John
; APPLICANT: Seeland, Sem
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Membrane Protein Genes;
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNA Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-111-470-8

Query Match
Best Local Similarity 74.4%; Score 29; DB 4; Length 238;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 31 EKPIRDLR 38

RESULT 12
US-07-803-622E-7
; Sequence 7, Application US/07803622E
; Patent No. 5523457
; GENERAL INFORMATION:
; APPLICANT: Keller, Walter
; APPLICANT: Langner, Joachim
; APPLICANT: Martin, Georges
; APPLICANT: Wahle, Elmar
; TITLE OF INVENTION: RECOMBINANT POLY(A) POLYMERASE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 611 West Sixth Street
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90017
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/803,622E
; FILING DATE: 27-NOV-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 195/296
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-489-1600
; TELEFAX: 213-955-0440
; TELEEX: 67-3510
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-07-803-622E-7

Query Match

Best Local Similarity 74.4%; Score 29; DB 1; Length 375;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8

Db 140 QEEVKDLR 147

RESULT 13

US-07-803-622E-9
Sequence 9, Application US/07803622E
Patent No. 5525497

GENERAL INFORMATION:

APPLICANT: Keller, Walter
APPLICANT: Lingner, JoachimAPPLICANT: Martin, Georges
APPLICANT: Mahle, ElmarTITLE OF INVENTION: RECOMBINANT POLY(A) POLYMERASE
NUMBER OF SEQUENCES: 9CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & LyonSTREET: 611 West Sixth Street
CITY: Los AngelesSTATE: CA
COUNTRY: USA

ZIP: 90017

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOSSOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/803,622EFILING DATE: 27-NOV-1991
CLASSIFICATION: 435ATTORNEY/AGENT INFORMATION:
NAME: Wardburg, Richard J.REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 195/296TELECOMMUNICATION INFORMATION:
TELEPHONE: 213-489-1600TELEFAX: 213-955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acidsTYPE: amino acid
TOPOLOGY: linearMOLECULE TYPE: protein
US-07-803-622E-9Query Match 74.4%; Score 29; DB 1; Length 375;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;

Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8

Db 140 QEEVKDLR 147

RESULT 14

US-08-985-908-24
Sequence 24, Application US/08985908
Patent No. 6004773

GENERAL INFORMATION:

APPLICANT: MASAYUKI ARAKI, MASARAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSHI NA
TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE
NUMBER OF SEQUENCES: 31CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTONCOUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatibleOPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,908FILING DATE: 05-DEC-1997
CLASSIFICATION: 435PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-325659FILING DATE: 05-DEC-1996
ATTORNEY/AGENT INFORMATION:NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 24:SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acidsTYPE: amino acid
TOPOLOGY: linearMOLECULE TYPE: protein
US-08-985-908-24Query Match 74.4%; Score 29; DB 3; Length 396;
Best Local Similarity 62.5%; Pred. No. 3.7e+02;

Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8

Db 308 EOEIRDLR 315

RESULT 15

US-08-650-275-2

Sequence 2, Application US/08650275
Patent No. 5788249

GENERAL INFORMATION:

APPLICANT: Braxton, Scott Michael
APPLICANT: Murry, Lynn E.TITLE OF INVENTION: HUMAN PROTEIN DISULFIDE ISOMERASE
NUMBER OF SEQUENCES: 35CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.STREET: 3174 Porter Drive
CITY: Palo AltoSTATE: CA
COUNTRY: U.S.

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM CompatibleOPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/650,275FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33,954REFERENCE/DOCKET NUMBER: PF-0067 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-852-0195TELEFAX: 415-855-0555
INFORMATION FOR SEQ ID NO: 2:SEQUENCE CHARACTERISTICS:
LENGTH: 406 amino acids

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: LUNGNOT04
; CLONE: 809200
US-08-650-275-2

Query Match 74.4%; Score 29; DB 1; Length 406;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 OEIRDL 7
| | | | |
Db 145 OEIRDL 150

Search completed: November 13, 2002, 13:18:09
Job time : 11.5745 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:18:18 : Search time 5.44681 Seconds
(without alignments)
22.121 Million cell updates/sec

Title: US-09-856-086-7

Perfect score: 39

Sequence: 1 EGERDLR 8

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB pep:*
- 2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB pep:*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB pep:*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB pep:*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*
- 6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB pep:*
- 7: /cgn2_6/ptodata/2/pubppaa/PCRUS_PUBCOMB pep:*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB pep:*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB pep:*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB pep:*
- 11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB pep:*
- 12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB pep:*
- 13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB pep:*
- 14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	32	82.1	183	10	US-09-815-242-13168 Sequence 13168, A
2	31	79.5	631	10	US-09-764-864-876 Sequence 876, App
3	30	76.9	618	10	US-09-925-300-1381 Sequence 1381, App
4	30	76.9	664	10	US-09-823-038A-48 Sequence 48, Appl
5	30	76.9	701	10	US-09-815-242-13411 Sequence 13411, A
6	30	76.9	1179	10	US-09-815-242-13608 Sequence 13608, A
7	29	74.4	65	10	US-09-815-242-12049 Sequence 12049, A
8	29	74.4	238	9	US-08-862-802-8 Sequence 8, Appl
9	29	74.4	398	10	US-09-815-242-13171 Sequence 13171, A
10	29	74.4	406	9	US-09-978-285A-309 Sequence 309, App
11	29	74.4	406	10	US-09-765-205-8 Sequence 8, Appl
12	29	74.4	446	10	US-09-866-572A-74 Sequence 74, Appl
13	29	74.4	466	10	US-09-923-779-155 Sequence 155, App
14	29	74.4	466	12	US-10-152-647-4 Sequence 3, Appl
15	29	74.4	466	12	US-10-152-647-4 Sequence 4, Appl
16	29	74.4	585	10	US-09-815-242-4973 Sequence 4973, Ap
17	29	74.4	585	10	US-09-815-242-10580 Sequence 10580, A
18	29	74.4	608	10	US-09-924-358-8 Sequence 8, Appl
19	29	74.4	1179	10	US-09-815-242-13262 Sequence 13262, A

20	29	74.4	3069	9	US-09-712-363-246 Sequence 246, App
21	28	71.8	18	10	US-09-918-063-66 Sequence 66, Appl
22	28	71.8	87	10	US-09-764-878-151 Sequence 151, App
23	28	71.8	146	10	US-09-939-880-349 Sequence 349, App
24	28	71.8	204	10	US-09-918-063-15 Sequence 15, Appl
25	28	71.8	204	10	US-09-918-063-30 Sequence 30, Appl
26	28	71.8	213	10	US-09-796-100-4 Sequence 4, Appl
27	28	71.8	227	10	US-09-918-063-74 Sequence 74, Appl
28	28	71.8	294	10	US-09-898-570-28 Sequence 28, Appl
29	28	71.8	294	10	US-09-898-570-30 Sequence 30, Appl
30	28	71.8	294	10	US-09-898-570-32 Sequence 32, Appl
31	28	71.8	326	10	US-09-918-063-3 Sequence 3, Appl
32	28	71.8	326	10	US-09-918-063-6 Sequence 6, Appl
33	28	71.8	329	10	US-09-918-063-18 Sequence 18, Appl
34	28	71.8	329	10	US-09-918-063-21 Sequence 21, Appl
35	28	71.8	337	10	US-09-925-300-1160 Sequence 1160, App
36	28	71.8	375	10	US-09-815-242-13383 Sequence 13383, A
37	28	71.8	375	10	US-09-815-242-13700 Sequence 13700, A
38	28	71.8	394	10	US-09-815-242-5574 Sequence 5574, Ap
39	28	71.8	394	10	US-09-815-242-12442 Sequence 12442, A
40	28	71.8	395	10	US-09-815-242-4887 Sequence 4887, Ap
41	28	71.8	395	10	US-09-815-242-10482 Sequence 10482, A
42	28	71.8	397	10	US-09-815-242-11995 Sequence 11995, A
43	28	71.8	397	10	US-09-815-242-12002 Sequence 12002, A
44	28	71.8	422	10	US-09-779-307-18 Sequence 18, Appl
45	28	71.8	422	10	US-09-779-307-20 Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-815-242-13168
Sequence 13168, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 1410
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13168
LENGTH: 183
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13168
Query Match 82.1% Score 32; DB 10; Length 183;
Best Local Similarity 85.7%; Pred. No. 11;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EOEIRDL 7
|||||:
Db 111 EOEIRDI 117

RESULT 2
US-09-764-864-876
; Sequence 876, Application US/09764864
; Patent No. US20020133753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 876
; LENGTH: 631
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-876

Query Match 79.5%; Score 31; DB 10; Length 631;
Best Local Similarity 75.0%; Pred. No. 66;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDL 8
|||||:
Db 149 EOEIRDLQ 156

RESULT 3
US-09-925-300-1381
; Sequence 1381, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1381
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (507)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (524)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (562)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1381

Query Match 76.9%; Score 30; DB 10; Length 618;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDL 8

Db 299 EOEIRDL 306
|:|:|:|

RESULT 4
US-09-823-038A-48
; Sequence 48, Application US/09823038A
; Patent No. US20020058335A1
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevlin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Morrison, Greg
; TITLE OF INVENTION: Compositions Isolated From Stromal Cells
; FILE REFERENCE: 11000.1037c3
; CURRENT APPLICATION NUMBER: US/09/823,038A
; CURRENT FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 664
; TYPE: PRT
; ORGANISM: Mouse
US-09-823-038A-48

Query Match 76.9%; Score 30; DB 10; Length 664;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDL 8
|:|:|:|
Db 604 DOEIRDLK 611

RESULT 5
US-09-815-242-13411
; Sequence 13411, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Treweek, John D.
; APPLICANT: Cair, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13411
; LENGTH: 701

TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13411

Query Match
Best Local Similarity 76.9%; Score 30; DB 10; Length 701;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDL 6
DB 666 EOEIRDL 671

RESULT 6
US-09-815-242-13608
Sequence 13608, Application US/09815242
Patent No. US2002061569A1

GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zykling, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13608
LENGTH: 1179
TYPE: PRT

ORGANISM: Streptococcus pneumoniae
US-09-815-242-13608

Query Match
Best Local Similarity 76.9%; Score 30; DB 10; Length 1179;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDL 7
DB 964 EOEIRDL 970

RESULT 7
US-09-815-242-12049
Sequence 12049, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zykling, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 12049
LENGTH: 65
TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12049

Query Match
Best Local Similarity 74.4%; Score 29; DB 10; Length 65;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDL 6
DB 55 EOEIRDL 60

RESULT 8
US-09-862-802-8
Sequence 8, Application US/09862802
Patent No. US20020165346A1
GENERAL INFORMATION:
APPLICANT: Schering-Plough Corporation
TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
FILE REFERENCE: SPO695B
CURRENT APPLICATION NUMBER: US/09/862,802
CURRENT FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 09/111,470
PRIOR FILING DATE: 1998-07-08
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 238
TYPE: PRT
ORGANISM: unknown
FEATURE:
OTHER INFORMATION: mammalian protein
US-09-862-802-8

Query Match
Best Local Similarity 74.4%; Score 29; DB 9; Length 238;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 31 EOEIRDLR 38

RESULT 9
US-09-815-242-13171
Sequence 13171, Application US/09815242
Patent No. US2002061569A1

```

: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zykand, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: TITLE OF INVENTION: Prokaryotes
: FILE REFERENCE: ELITRA.011a
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-05-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 13171
: LENGTH: 398
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
US-09-815-242-13171

Query Match          74.4%; Score 29; DB 10; Length 398;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDL 7
Db 154 EMEIRDL 160

RESULT 10
US-09-978-295A-309
: Sequence 309, Application US/09/8295A
: Patent No. US20020156006A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Geo, Wei-Olang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kijavala, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
```

```

: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630PIC11
: CURRENT APPLICATION NUMBER: US/09/978,295A
: CURRENT FILING DATE: 2001-10-15
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: PRIOR APPLICATION NUMBER: 60/078004
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: 60/078886
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078936
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078939
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079664
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079663
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079786
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079920
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/079923
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/080105
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080107
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080165
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080194
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080327
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080328
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080333
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080334
: PRIOR FILING DATE: 1998-04-01
```

PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598

PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 74.4%; Score 29; DB 9; Length 406;
Best Local Similarity 100.0%; Pred. No. 1e+02; 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

OY 2 OEIRDL 7
DB 145 OEIRDL 150

RESULT 11
US-09-765-205-8
Sequence 8, Application US/09765205
Patent No. US20020034800A1
GENERAL INFORMATION:
APPLICANT: Caeo, LI
TITLE OF INVENTION: BONE MARROW SECRETED PROTEINS AND POLYNUCLEOTIDES
FILE-REFERENCE: 1458 004/200130 449
CURRENT APPLICATION NUMBER: US/09/765,205
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: US/09/212,440
PRIOR FILING DATE: 1998-12-16
NUMBER OF SEQ ID NOS: 46
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 406
TYPE: PRT
ORGANISM: human
US-09-765-205-8

Query Match 74.4%; Score 29; DB 10; Length 406;
Best Local Similarity 100.0%; Pred. No. 1e+02; 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

OY 2 OEIRDL 7
DB 145 OEIRDL 150

RESULT 12
US-09-866-572A-74
Sequence 74, Application US/09866572A
Patent No. US20020138659A1
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney et al.

;; TITLE OF INVENTION: Transacylases of the Facilitaxel Biosynthetic Pathway
;; FILE REFERENCE: 53679
;; CURRENT APPLICATION NUMBER: US/09/866,572A
;; CURRENT FILING DATE: 2001-05-25
;; PRIOR APPLICATION NUMBER: 09/457,046
;; PRIOR FILING DATE: 1999-12-07
;; NUMBER OF SEQ ID NOS: 74
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 74
;; LENGTH: 446
;; TYPE: PRT
;; ORGANISM: Arabidopsis thaliana
US-09-866-572A-74

Query Match 74.4%; Score 29; DB 10; Length 446;
Best Local Similarity 62.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 EOERDLR 8
I:|:|:|
Db 315 EEAVDLR 322

RESULT 13
US-09-923-779-155
;; Sequence 155, Application US/09923779
;; Patent No. US20020076721A1
;; GENERAL INFORMATION:
;; APPLICANT: Pyle, Ruth A.
;; APPLICANT: Xu, Jiangchun
;; APPLICANT: Kalos, Michael D.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
;; FILE REFERENCE: 210121.553
;; CURRENT APPLICATION NUMBER: US/09/923,779
;; CURRENT FILING DATE: 2001-08-06
;; NUMBER OF SEQ ID NOS: 155
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 155
;; LENGTH: 466
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-923-779-155

Query Match 74.4%; Score 29; DB 10; Length 466;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOERDLR 8
I:|:|:|
Db 151 EEEMREL 158

RESULT 14
US-10-152-647-3
;; Sequence 3, Application US/10152647
;; Patent No. US20020137110A1
;; GENERAL INFORMATION:
;; APPLICANT: MORISHIMA, No. US20020137110A1uhltro,
;; APPLICANT: NAKANISHI, Keiko,
;; APPLICANT: SHIBATA, Takehiko
;; TITLE OF INVENTION: Antibody against cleavage product of vimentin
;; FILE REFERENCE: 522.1004
;; CURRENT APPLICATION NUMBER: US/10/152,647
;; CURRENT FILING DATE: 2002-05-21
;; PRIOR APPLICATION NUMBER: US/09/610,401
;; PRIOR FILING DATE: 2000-07-05
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 3
;; LENGTH: 466
;; TYPE: PRT
;; ORGANISM: Homo sapiens

US-10-152-647-3

Query Match 74.4%; Score 29; DB 12; Length 466;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOERDLR 8
I:|:|:|
Db 151 EEEMREL 158

RESULT 15
US-10-152-647-4
;; Sequence 4, Application US/10152647
;; Patent No. US20020137110A1
;; GENERAL INFORMATION:
;; APPLICANT: MORISHIMA, No. US20020137110A1uhltro,
;; APPLICANT: NAKANISHI, Keiko,
;; APPLICANT: SHIBATA, Takehiko
;; TITLE OF INVENTION: Antibody against cleavage product of vimentin
;; FILE REFERENCE: 522.1004
;; CURRENT APPLICATION NUMBER: US/10/152,647
;; CURRENT FILING DATE: 2002-05-21
;; PRIOR APPLICATION NUMBER: US/09/610,401
;; PRIOR FILING DATE: 2000-07-05
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 4
;; LENGTH: 466
;; TYPE: PRT
;; ORGANISM: Mus sp.
US-10-152-647-4

Query Match 74.4%; Score 29; DB 12; Length 466;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOERDLR 8
I:|:|:|
Db 151 EEEMREL 158

Search completed: November 13, 2002, 13:40:20
Job time : 6.58967 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:16:18 ; Search time 160.511 Seconds
(without alignments)
32.134 Million cell updates/sec

Title: US-09-856-086-7

Perfect score: 39

Sequence: 1 EOEIRDLR 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:

1: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US110_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US111_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	39	100.0	542 21 US-09-791-537-92908	Sequence 92908, A
2	39	100.0	542 21 US-09-791-537-125497	Sequence 125497, A
3	39	100.0	542 26 US-10-203-331-57	Sequence 57, Appl
4	39	100.0	543 20 US-09-611-526-3132	Sequence 3132, Ap
5	39	100.0	543 21 US-09-791-537-34475	Sequence 34475, A
6	39	100.0	543 21 US-09-791-537-56903	Sequence 56903, A

7	39	100.0	544 21	US-09-791-537-86802	Sequence 86802, A
8 <td>39</td> <td>100.0</td> <td>548 21</td> <td>US-09-791-537-125495</td> <td>Sequence 125495, A</td>	39	100.0	548 21	US-09-791-537-125495	Sequence 125495, A
9 <td>39</td> <td>100.0</td> <td>554 21</td> <td>US-09-791-537-94465</td> <td>Sequence 94465, A</td>	39	100.0	554 21	US-09-791-537-94465	Sequence 94465, A
10 <td>39</td> <td>100.0</td> <td>555 21</td> <td>US-09-791-537-18635</td> <td>Sequence 18635, A</td>	39	100.0	555 21	US-09-791-537-18635	Sequence 18635, A
11 <td>39</td> <td>92.3</td> <td>229 27</td> <td>US-60-207-216-737</td> <td>Sequence 737, Appl</td>	39	92.3	229 27	US-60-207-216-737	Sequence 737, Appl
12 <td>36</td> <td>92.3</td> <td>404 21</td> <td>US-09-760-475-2945</td> <td>Sequence 2945, Ap</td>	36	92.3	404 21	US-09-760-475-2945	Sequence 2945, Ap
13 <td>36</td> <td>92.3</td> <td>455 21</td> <td>US-09-791-537-32570</td> <td>Sequence 32570, A</td>	36	92.3	455 21	US-09-791-537-32570	Sequence 32570, A
14 <td>36</td> <td>92.3</td> <td>608 1</td> <td>PCT-US02-31096-25</td> <td>Sequence 25, Appl</td>	36	92.3	608 1	PCT-US02-31096-25	Sequence 25, Appl
15 <td>35</td> <td>89.7</td> <td>543 21</td> <td>US-09-791-537-46537</td> <td>Sequence 46537, A</td>	35	89.7	543 21	US-09-791-537-46537	Sequence 46537, A
16 <td>35</td> <td>89.7</td> <td>543 21</td> <td>US-09-791-537-49881</td> <td>Sequence 49881, A</td>	35	89.7	543 21	US-09-791-537-49881	Sequence 49881, A
17 <td>35</td> <td>89.7</td> <td>556 21</td> <td>US-09-791-537-30712</td> <td>Sequence 30712, A</td>	35	89.7	556 21	US-09-791-537-30712	Sequence 30712, A
18 <td>35</td> <td>89.7</td> <td>556 21</td> <td>US-09-791-537-120110</td> <td>Sequence 120110, A</td>	35	89.7	556 21	US-09-791-537-120110	Sequence 120110, A
19 <td>35</td> <td>89.7</td> <td>556 21</td> <td>US-09-791-537-120823</td> <td>Sequence 120823, A</td>	35	89.7	556 21	US-09-791-537-120823	Sequence 120823, A
20 <td>35</td> <td>89.7</td> <td>574 21</td> <td>US-09-791-537-26592</td> <td>Sequence 26592, A</td>	35	89.7	574 21	US-09-791-537-26592	Sequence 26592, A
21 <td>35</td> <td>89.7</td> <td>574 21</td> <td>US-09-791-537-46529</td> <td>Sequence 46529, A</td>	35	89.7	574 21	US-09-791-537-46529	Sequence 46529, A
22 <td>35</td> <td>89.7</td> <td>574 21</td> <td>US-09-791-537-73376</td> <td>Sequence 73376, A</td>	35	89.7	574 21	US-09-791-537-73376	Sequence 73376, A
23 <td>35</td> <td>89.7</td> <td>587 21</td> <td>US-09-791-537-120809</td> <td>Sequence 120809, A</td>	35	89.7	587 21	US-09-791-537-120809	Sequence 120809, A
24 <td>35</td> <td>89.7</td> <td>599 21</td> <td>US-09-791-537-17426</td> <td>Sequence 17426, A</td>	35	89.7	599 21	US-09-791-537-17426	Sequence 17426, A
25 <td>34</td> <td>87.2</td> <td>633 25</td> <td>US-10-153-881-20594</td> <td>Sequence 20594, A</td>	34	87.2	633 25	US-10-153-881-20594	Sequence 20594, A
26 <td>34</td> <td>87.2</td> <td>633 27</td> <td>US-60-207-316-110</td> <td>Sequence 110, Appl</td>	34	87.2	633 27	US-60-207-316-110	Sequence 110, Appl
27 <td>34</td> <td>87.2</td> <td>638 21</td> <td>US-09-791-537-36553</td> <td>Sequence 36553, A</td>	34	87.2	638 21	US-09-791-537-36553	Sequence 36553, A
28 <td>34</td> <td>87.2</td> <td>747 21</td> <td>US-09-791-537-73306</td> <td>Sequence 73306, A</td>	34	87.2	747 21	US-09-791-537-73306	Sequence 73306, A
29 <td>34</td> <td>87.2</td> <td>907 21</td> <td>US-09-791-537-28076</td> <td>Sequence 28076, A</td>	34	87.2	907 21	US-09-791-537-28076	Sequence 28076, A
30 <td>34</td> <td>87.2</td> <td>919 21</td> <td>US-09-791-537-28079</td> <td>Sequence 28079, A</td>	34	87.2	919 21	US-09-791-537-28079	Sequence 28079, A
31 <td>34</td> <td>87.2</td> <td>1011 21</td> <td>US-09-791-537-28058</td> <td>Sequence 28058, A</td>	34	87.2	1011 21	US-09-791-537-28058	Sequence 28058, A
32 <td>34</td> <td>87.2</td> <td>1022 21</td> <td>US-09-791-537-4782</td> <td>Sequence 4782, Ap</td>	34	87.2	1022 21	US-09-791-537-4782	Sequence 4782, Ap
33 <td>34</td> <td>87.2</td> <td>1022 27</td> <td>US-60-406-385-230</td> <td>Sequence 230, Appl</td>	34	87.2	1022 27	US-60-406-385-230	Sequence 230, Appl
34 <td>34</td> <td>87.2</td> <td>1023 27</td> <td>US-09-791-537-97590</td> <td>Sequence 97590, A</td>	34	87.2	1023 27	US-09-791-537-97590	Sequence 97590, A
35 <td>34</td> <td>87.2</td> <td>1093 21</td> <td>US-60-203-511-1</td> <td>Sequence 1, Appl1</td>	34	87.2	1093 21	US-60-203-511-1	Sequence 1, Appl1
36 <td>34</td> <td>87.2</td> <td>1111 27</td> <td>US-60-248-923-71</td> <td>Sequence 71, Appl</td>	34	87.2	1111 27	US-60-248-923-71	Sequence 71, Appl
37 <td>34</td> <td>87.2</td> <td>1975 20</td> <td>US-09-614-150-13074</td> <td>Sequence 13074, A</td>	34	87.2	1975 20	US-09-614-150-13074	Sequence 13074, A
38 <td>34</td> <td>87.2</td> <td>1975 27</td> <td>US-60-167-217-13120</td> <td>Sequence 13120, A</td>	34	87.2	1975 27	US-60-167-217-13120	Sequence 13120, A
39 <td>34</td> <td>87.2</td> <td>1975 27</td> <td>US-60-173-464-10670</td> <td>Sequence 10670, A</td>	34	87.2	1975 27	US-60-173-464-10670	Sequence 10670, A
40 <td>34</td> <td>87.2</td> <td>1975 27</td> <td>US-60-191-637-13115</td> <td>Sequence 13115, A</td>	34	87.2	1975 27	US-60-191-637-13115	Sequence 13115, A
41 <td>34</td> <td>87.2</td> <td>1975 27</td> <td>US-60-191-681-10316</td> <td>Sequence 10316, A</td>	34	87.2	1975 27	US-60-191-681-10316	Sequence 10316, A
42 <td>34</td> <td>87.2</td> <td>2080 27</td> <td>US-60-161-932-2022</td> <td>Sequence 2022, Ap</td>	34	87.2	2080 27	US-60-161-932-2022	Sequence 2022, Ap
43 <td>34</td> <td>84.6</td> <td>85 18</td> <td>US-09-462-917A-111</td> <td>Sequence 111, Appl</td>	34	84.6	85 18	US-09-462-917A-111	Sequence 111, Appl
44 <td>33</td> <td>84.6</td> <td>85 18</td> <td>US-09-462-917A-114</td> <td>Sequence 114, Appl</td>	33	84.6	85 18	US-09-462-917A-114	Sequence 114, Appl
45 <td>33</td> <td>84.6</td> <td>244 16</td> <td>US-09-248-796-15505</td> <td>Sequence 15505, A</td>	33	84.6	244 16	US-09-248-796-15505	Sequence 15505, A

ALIGNMENTS

RESULT 1
US-09-791-537-92908
Sequence 92908, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Biomedix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: DANZER, JOSEPH
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 92908
LENGTH: 542
TYPE: PRT
ORGANISM: Mus musculus
US-09-791-537-92908

Query Match 100.0%; Score 39; DB 21; Length 542;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
Db 140 EOEIRDLR 147

```
RESULT 2
US-09-791-537-125497
; Sequence 125497, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 125497
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-125497

Query Match          100.0%; Score 39; DB 21; Length 542;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 140 EOEIRDLR 147

RESULT 3
US-10-205-331-57
; Sequence 57, Application US/10205331
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alstair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Plinock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: NF-L
US-10-205-331-57

Query Match          100.0%; Score 39; DB 26; Length 542;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 140 EOEIRDLR 147

RESULT 4
US-09-611-526-3132
; Sequence 3132, Application US/09611526
; GENERAL INFORMATION:
; APPLICANT: OYA, TOSHIO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: KAWAI, YURI
US-09-611-526-3132

APPLICANT: WAKAMATSU, AI
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: NAGAI, KEIICHI
APPLICANT: KOJIMA, SHINICHI
APPLICANT: OTSUKI, TERSUJI
APPLICANT: KOGA, HISASHI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH CDNA8
; FILE REFERENCE: 08335/0122
; CURRENT APPLICATION NUMBER: US/09/611,526
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: JP 1999-194486
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: JP 2000-118774
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183765
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 4484
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3132
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-611-526-3132

Query Match          100.0%; Score 39; DB 20; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 140 EOEIRDLR 147

RESULT 5
US-09-791-537-34475
; Sequence 34475, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY M
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34475
; LENGTH: 543
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-34475

Query Match          100.0%; Score 39; DB 21; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 140 EOEIRDLR 147

RESULT 6
US-09-791-537-56903
; Sequence 56903, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY M
; FILE REFERENCE: 261/210
US-09-791-537-56903
```

;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 56903
;; LENGTH: 543
;; TYPE: PRF
;; ORGANISM: Mus musculus
US-09-791-537-56903

Query Match
Best Local Similarity 100.0%; Score 39; DB 21; Length 543;
Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 140 EOEIRDLR 147

RESULT 7
US-09-791-537-86802
;; Sequence 86802, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biocomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 86802
;; LENGTH: 544
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-09-791-537-86802

Query Match
Best Local Similarity 100.0%; Score 39; DB 21; Length 544;
Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 139 EOEIRDLR 146

RESULT 8
US-09-791-537-125495
;; Sequence 125495, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biocomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 125495
;; LENGTH: 548
;; TYPE: PRF
;; ORGANISM: Sus scrofa
US-09-791-537-125495

Query Match
Best Local Similarity 100.0%; Score 39; DB 21; Length 548;
Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8

DB 138 EOEIRDLR 145

RESULT 9
US-09-791-537-94465
;; Sequence 94465, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biocomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 94465
;; LENGTH: 554
;; TYPE: PRF
;; ORGANISM: Bos taurus
US-09-791-537-94465

Query Match
Best Local Similarity 100.0%; Score 39; DB 21; Length 554;
Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 139 EOEIRDLR 146

RESULT 10
US-09-791-537-18635
;; Sequence 18635, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biocomix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 18635
;; LENGTH: 555
;; TYPE: PRF
;; ORGANISM: Bos taurus
US-09-791-537-18635

Query Match
Best Local Similarity 100.0%; Score 39; DB 21; Length 555;
Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 140 EOEIRDLR 147

RESULT 11
US-60-207-216-737
;; Sequence 737, Application US/60207216
;; GENERAL INFORMATION:
;; APPLICANT: Beasley, Ellen
;; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
;; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
;; FILE REFERENCE: CLO00580
;; CURRENT APPLICATION NUMBER: US/60/207, 216
;; CURRENT FILING DATE: 2000-05-26

```

? NUMBER OF SEQ ID NOS: 898
? SOFTWARE: FaalSeq for Windows Version 4.0
? SEQ ID NO 737
? LENGTH: 229
? TYPE: PRT
? ORGANISM: HUMAN
? FEATURE:
? NAME/KEY: VARIANT
? LOCATION: (1)...(229)
? OTHER INFORMATION: xaa - Any Amino Acid
? OS: 60-207-216-737

```

Query Match	92.3%	Score 36:	DB 27:	Length 229:
Best Local Similarity	75.0%	Pred. No.	1.7e+02:	
Matches	6:	Conservative	2:	Mismatches 0:
				Indels 0:
				Gaps 0:

QY	1	EQEIRDLR	8
		: :	
Db	193	EQEVRDIR	200

```

RESULT 12
US-09-760-475-2945
; Sequence 2945, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT49
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2945
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-475-2945

```

Query Match	92.3%	Score 36:	DB 21:	Length 404:
Best Local Similarity	75.0%	Pred. NO.	3.6e+02:	
Matches	6:	Conservative	2:	Mismatches 0; Gaps 0;

QY	1	EQEIRDLR	8
		: :	
Db	368	EQEVRDIR	375

```

RESULT 13
US-09-791-537-32570
/ Sequence 32570, Application US/09791537
/ GENERAL INFORMATION:
/ APPLICANT: Biomomix, Inc.
/ APPLICANT: Debe, Derek
/ TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
/ TITLE OF INVENTION: METHODS OF USE THEREOF
/ FILE REFERENCE: 261/210
/ CURRENT APPLICATION NUMBER: US/09/791,537
/ CURRENT FILING DATE: 2001-02-22
/ NUMBER OF SEQ ID NOS: 153055
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 32570
/ LENGTH: 455
/ TYPE: prt
/ ORGANISM: Scyllorhinus stellaris
US-09-791-537-32570

```

Query Match	92.38;	Score 36;	DB 21;	Length 455;
Best Local Similarity	87.38;	Pred. No. 4.2e+02;		
Matches 7; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 EQEIRDLR 8

```

Db      140 EQEMRDLR 147
          |||:||||

```

RESULT 14
PCT-US02-31096-25
Sequence 25, Application PC/TUS0231096
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: YANG, Junming
APPLICANT: LU, Dyrong, Alina M.
APPLICANT: YDE, Henry
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: WARREN, Bridget A.
APPLICANT: DUGGAN, Brendan M.
APPLICANT: FORSYTHE, Ian J.
APPLICANT: LEE, Ernestine A.
APPLICANT: HAFALIA, April J. A.
APPLICANT: RAMKOWAR, Jayalaxmi
APPLICANT: CHAWLA, Nalinder K.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: BECHA, Shanya D.
APPLICANT: GORVAD, Ann E.
APPLICANT: TRAN, Uyem K.
APPLICANT: LI, Joane X.
APPLICANT: YAO, Monique G.
APPLICANT: ISON, Craig H.
APPLICANT: GRIFFIN, Jennifer A.
APPLICANT: LEE, Soo Yeun
APPLICANT: CHANG, Hsin-Ru
APPLICANT: EMERLING, Brooke M.
APPLICANT: TANG, Y. Tom
APPLICANT: LAU, Freeli G.
APPLICANT: KABLE, Amy E.
APPLICANT: MARQUIS, Joseph P.
APPLICANT: JIANG, Xin
APPLICANT: JACKSON, Alan A.
APPLICANT: ZEBARADIAN, Yeganeh
APPLICANT: SHARNAKAR, Arita
APPLICANT: WILSON, Amy D.
APPLICANT: JIN, Pei
APPLICANT: RICHARDSON, Thomas W.
APPLICANT: BHATTI, Umesh
APPLICANT: BURRILL, John D.
APPLICANT: LEE, Sally
APPLICANT: BLAKE, Julie J.
APPLICANT: HO, Anne
APPLICANT: ZHENG, Wenjin
APPLICANT: GAO, Jing
TITLE OF INVENTION: ENZYMES
FILE REFERENCE: PF-1214 PCT
CURRENT APPLICATION NUMBER: PCT/US02/31096
PRIOR FILING DATE: 2002-09-26
PRIOR APPLICATION NUMBER: US 60/326,388
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 60/328,979
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/346,034
PRIOR FILING DATE: 2001-10-19
PRIOR APPLICATION NUMBER: US 60/348,284
PRIOR FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: US 60/338,048
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: US 60/332,340
PRIOR FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: US 60/340,357
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/387,119
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: US 60/368,799
PRIOR FILING DATE: 2002-03-29
PRIOR APPLICATION NUMBER: US 60/368,722
PRIOR FILING DATE: 2002-03-29

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 106
SOFTWARE: PERL Program
SEQ ID NO 25
LENGTH: 608
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID NO: 7503416CD1
PCT-US02-31096-25

Query Match 92.3%; Score 36; DB 1; Length 608;
Best Local Similarity 75.0%; Pred. No. 6e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
|||||

DB 572 EOEVRDIR 579

RESULT 15

US-09-791-537-46537
Sequence 46537, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Bionomix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB

FILE REFERENCE: 261/210

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 46537

LENGTH: 543

TYPE: PRT

ORGANISM: Homo sapiens

US-09-791-537-46537

Query Match 89.7%; Score 35; DB 21; Length 543;
Best Local Similarity 75.0%; Pred. No. 8.2e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
|||||

DB 235 EODVRDLR 242

Search completed: November 13, 2002, 13:39:21
Job time : 161.511 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:17:03 ; Search time 3.23404 Seconds
(without alignments)
29.431 Million cell updates/sec

Title: US-09-856-086-7

Perfect score: 39

Sequence: 1 EOEIRDLR 8

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 41632 seqs, 11897606 residues

Total number of hits satisfying chosen parameters: 41632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/US06_PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	8	5	US-09-856-086-7
2	30	76.9	143	5	US-09-513-999C-7973
3	29	74.4	6	5	US-09-856-086-6
4	29	74.4	116	1	PCT-US02-31996-16
5	29	74.4	236	1	PCT-US02-31996-12
6	29	74.4	406	6	US-10-145-087A-309
7	29	74.4	406	6	US-10-143-031A-309
8	29	74.4	406	6	US-10-145-092A-309
9	29	74.4	406	6	US-10-162-522A-309
10	29	74.4	406	6	US-10-165-038A-309
11	29	74.4	406	6	US-10-165-353-309
12	29	74.4	406	6	US-10-170-481A-309
13	29	74.4	406	6	US-10-172-039A-309
14	29	74.4	406	6	US-10-145-016A-309
15	29	74.4	406	6	US-10-145-088A-309
16	29	74.4	406	6	US-10-145-129A-309
17	29	74.4	406	6	US-10-131-813A-430
18	29	74.4	406	6	US-10-131-819A-430
19	29	74.4	406	6	US-10-131-823A-430
20	29	74.4	406	6	US-10-131-824A-430
21	29	74.4	406	6	US-10-131-826A-430
22	29	74.4	406	6	US-10-131-829A-430
23	29	74.4	406	6	US-10-125-826A-430
24	29	74.4	406	6	US-10-127-829A-430
25	29	74.4	406	6	US-10-127-831A-430
26	29	74.4	406	6	US-10-127-835A-430

27	29	74.4	406	6	US-10-127-837A-430	Sequence 430, App
28	29	74.4	406	6	US-10-127-842A-430	Sequence 430, App
29	29	74.4	406	6	US-10-127-850A-430	Sequence 430, App
30	29	74.4	406	6	US-10-127-901A-430	Sequence 430, App
31	29	74.4	406	6	US-10-128-689A-430	Sequence 430, App
32	29	74.4	406	6	US-10-131-830A-430	Sequence 430, App
33	29	74.4	406	6	US-10-131-833A-430	Sequence 430, App
34	29	74.4	406	6	US-10-131-837A-430	Sequence 430, App
35	29	74.4	406	6	US-10-125-930A-430	Sequence 430, App
36	29	74.4	406	6	US-10-127-825A-430	Sequence 430, App
37	29	74.4	406	6	US-10-127-838B-430	Sequence 430, App
38	29	74.4	406	6	US-10-127-843A-430	Sequence 430, App
39	29	74.4	406	6	US-10-127-849A-430	Sequence 430, App
40	29	74.4	406	6	US-10-128-684A-430	Sequence 430, App
41	29	74.4	406	6	US-10-128-685A-430	Sequence 430, App
42	29	74.4	406	6	US-10-128-686A-430	Sequence 430, App
43	29	74.4	406	6	US-10-128-690A-430	Sequence 430, App
44	29	74.4	406	6	US-10-128-693A-430	Sequence 430, App
45	29	74.4	406	6	US-10-131-821A-430	Sequence 430, App

ALIGNMENTS

RESULT 1
US-09-856-086-7
Sequence 7, Application US/09856086
GENERAL INFORMATION:
APPLICANT: EBRINGER, ALAN
TITLE OF INVENTION: DIAGNOSIS OF DEMYELINATING OR SPONGIFORM DISEASE
FILE REFERENCE: 78104.040
CURRENT APPLICATION NUMBER: US/09/856, 086
CURRENT FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 8
TYPE: PRT
ORGANISM: Homo sapiens, Bos
US-09-856-086-7

Query Match 100.0%; Score 39; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.8e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
Db 1 EOEIRDLR 8

RESULT 2
US-09-513-999C-7973
Sequence 7973, Application US/09513999C

GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclelet, A.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

FILE REFERENCE: 59, US2, REG
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 7973

LENGTH: 143
TYPE: PRT
ORGANISM: Homo sapiens

US-09-513-999C-7973

Query Match 76.9%; Score 30; DB 5; Length 143;
Best Local Similarity 75.0%; Pred. No. 13;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 EOEIRDLR 8
11111111
DB 99 EOEVRRLR 106

RESULT 3

US-09-856-086-6
Sequence 6: Application US/09856086
GENERAL INFORMATION:
APPLICANT: EBBINGER, ALAN
TITLE OF INVENTION: DIAGNOSIS OF DEMENTELINATING OR SPONGIFORM DISEASE
FILE REFERENCE: 78104.040
CURRENT APPLICATION NUMBER: US/09/856,086
CURRENT FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 6
TYPE: PRT
ORGANISM: Homo sapiens, Bos
US-09-856-086-6

Query Match 74.4%; Score 29; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.8e+04; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 3 EIRDLR 8
11111111
DB 1 EIRDLR 6

RESULT 4

PCT-US02-31996-16
Sequence 16: Application PC/TUS0231996
GENERAL INFORMATION:
APPLICANT: IMMUNEX CORPORATION
APPLICANT: Butz, Eric A.
TITLE OF INVENTION: MAMMALIAN C-TYPE LECTINS
FILE REFERENCE: 3318-WO
CURRENT APPLICATION NUMBER: PCT/US02/31996
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: US 60/328,026
PRIOR FILING DATE: 2001-10-09
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 116
TYPE: PRT
ORGANISM: Mus sp.
PCT-US02-31996-16

Query Match 74.4%; Score 29; DB 1; Length 116;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
11111111
DB 31 EIRDLR 38

RESULT 5

PCT-US02-31996-12
Sequence 12: Application PC/TUS0231996
GENERAL INFORMATION:
APPLICANT: IMMUNEX CORPORATION
APPLICANT: Butz, Eric A.
TITLE OF INVENTION: MAMMALIAN C-TYPE LECTINS
FILE REFERENCE: 3318-WO
CURRENT APPLICATION NUMBER: PCT/US02/31996

CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: US 60/328,026
PRIOR FILING DATE: 2001-10-09
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 236
TYPE: PRT
ORGANISM: Mus sp.
PCT-US02-31996-12

Query Match 74.4%; Score 29; DB 1; Length 236;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
11111111
DB 31 EIRDLR 38

RESULT 6

US-10-145-087A-309
Sequence 309: Application US/10145087A
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gunney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tunas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C47
CURRENT APPLICATION NUMBER: US/10/145,087A
CURRENT FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649

;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; Remaining prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 624
;; SEQ ID NO 309
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-145-087A-309

Query Match 74.4%; Score 29; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OEIRDL 7
DB 145 OEIRDL 150

RESULT 7
US-10-143-031A-309
; Sequence 309, Application US/10143031A
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC39
; CURRENT APPLICATION NUMBER: US/10/143,031A
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649

;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; Remaining prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 624
;; SEQ ID NO 309
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-143-031A-309

Query Match 74.4%; Score 29; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OEIRDL 7
DB 145 OEIRDL 150

RESULT 8
US-10-145-092A-309
; Sequence 309, Application US/10145092A
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC45
; CURRENT APPLICATION NUMBER: US/10/145,092A
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649

```
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; Remaining Prior Application data removed - See file wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 624
;; SEQ ID NO 309
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-145-092A-309
```

```
Query Match          74.4%; Score 29; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 QEIRDL 7
        |||||
Db      145 QEIRDL 150
```

```
RESULT 9
US-10-162-522A-309
; Sequence 309, Application US/10162522A
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC56
; CURRENT APPLICATION NUMBER: US/10/162,522A.
; PRIOR APPLICATION NUMBER: 2002-10-10.
; PRIOR FILING DATE: 1997-11-03.
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
```

```
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; Remaining Prior Application data removed - See file wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 624
;; SEQ ID NO 309
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-162-522A-309
```

```
Query Match          74.4%; Score 29; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2 QEIRDL 7
        |||||
Db      145 QEIRDL 150
```

```
RESULT 10
US-10-165-038A-309
; Sequence 309, Application US/10165038A
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC29
; CURRENT APPLICATION NUMBER: US/10/165,038A
; PRIOR APPLICATION NUMBER: 2002-10-10
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
```

```
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 624
;; SEQ ID NO 309
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-165-038A-309

Query Match          74.4%; Score 29; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OEIRDL 7
DB 145 OEIRDL 150

RESULT 11
US-10-165-353-309
;; Sequence 309, Application US/10165353
;; GENERAL INFORMATION:
;; APPLICANT: Ashkenazi, Avi
;; APPLICANT: Baker Kevin P.
;; APPLICANT: Botstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan
;; APPLICANT: Ferrara, Napoleon
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Fond, Sherman
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerber, Hanspeter
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Hillan, Kenneth J.
;; APPLICANT: Kijavlin, Ivar J.
;; APPLICANT: Kuo, Sophia S.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Shelton, David L.
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC40
CURRENT APPLICATION NUMBER: US/10/165,353
PRIOR FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
```

```
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 624
;; SEQ ID NO 309
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-165-353-309

Query Match          74.4%; Score 29; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OEIRDL 7
DB 145 OEIRDL 150

RESULT 12
US-10-170-481A-309
;; Sequence 309, Application US/10170481A
;; GENERAL INFORMATION:
;; APPLICANT: Ashkenazi, Avi
;; APPLICANT: Baker Kevin P.
;; APPLICANT: Botstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan
;; APPLICANT: Ferrara, Napoleon
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Fond, Sherman
;; APPLICANT: Gao, Wei-Qiang
;; APPLICANT: Gerber, Hanspeter
;; APPLICANT: Gerritsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Hillan, Kenneth J.
;; APPLICANT: Kijavlin, Ivar J.
;; APPLICANT: Kuo, Sophia S.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Shelton, David L.
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC53
CURRENT APPLICATION NUMBER: US/10/170,481A
PRIOR FILING DATE: 2002-10-10
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
```

```
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 624
;; SEQ ID NO 309
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-170-481A-309

Query Match          74.4%; Score 29; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 QEI RDL 7
DB 145 QEI RDL 150
```

```
RESULT 13
US-10-172-039A-309
; Sequence 309, Application US/10172039A
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Geo, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C30
; CURRENT APPLICATION NUMBER: US/10/172.039A
; PRIOR FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
```

```
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 624
;; SEQ ID NO 309
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-172-039A-309

Query Match          74.4%; Score 29; DB 6; Length 406;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2 QEI RDL 7
DB 145 QEI RDL 150
```

```
RESULT 14
US-10-145-016A-309
; Sequence 309, Application US/10145016A
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Geo, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C52
; CURRENT APPLICATION NUMBER: US/10/145.016A
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
```

;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 624
;; SEQ ID NO 309
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-145-016A-309

Query Match 74.4%: Score 29; DB 6; Length 406;
Best Local Similarity 100.0%: Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OEIRDL 7
|||||
DB 145 OEIRDL 150

RESULT 15
US-10-145-088A-309
; Sequence 309, Application US/10145088A
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gunney, Austin L.
; APPLICANT: Hillen, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC49
; CURRENT APPLICATION NUMBER: US/10/145,088A
; PRIOR APPLICATION NUMBER: 2002-10-10
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649

;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 624
;; SEQ ID NO 309
;; LENGTH: 406
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-145-088A-309

Query Match 74.4%: Score 29; DB 6; Length 406;
Best Local Similarity 100.0%: Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 OEIRDL 7
|||||
DB 145 OEIRDL 150

Search completed: November 13, 2002, 13:39:41
Job time : 4.23404 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 : Search time 12.4255 seconds
(without alignments)
61.895 Million cell updates/sec

Title: US-09-856-086-7

Perfect score: 39

Sequence: 1 EQRDLR 8

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%

Listing first 45 summaries

Database:

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	100.0	543	1 QPM5L	neurofilament trip
2	39	100.0	544	1 S07144	neurofilament trip
3	39	100.0	548	1 QPGL	neurofilament trip
4	39	100.0	554	1 JMO094	neurofilament trip
5	36	92.3	221	2 A40785	hypothetical prote
6	36	92.3	440	2 S69546	phosphoenolpyruvat
7	35	89.7	430	1 VEMSGF	glial fibrillary a
8	35	89.7	469	2 A44841	low molecularweig
9	35	89.7	543	1 KIHUPL	pyruvate kinase (E
10	35	89.7	556	2 B46024	neurofilament-L su
11	35	89.7	574	1 KIHUPL	pyruvate kinase (E
12	34	87.2	231	2 H83875	two-component resp
13	34	87.2	497	2 T51195	hypothetical prote
14	34	87.2	1023	2 B38932	phospholipase C (E
15	34	87.2	1176	1 A48047	phospholipase C (E
16	33	84.6	415	2 T00140	capsid protein - S
17	33	84.6	458	2 S09228	intermediate filam
18	33	84.6	464	2 A56600	anti-silencing pro
19	33	84.6	525	2 S30767	lamin - chicken
20	33	84.6	657	2 S05517	neurofilament trip
21	33	84.6	818	2 S15762	probable lipoxigen
22	33	84.6	914	2 T29999	hypothetical prote
23	33	84.6	2288	2 T29999	conserved hypothet
24	32	82.1	105	2 E75442	gp46 protein - Myc
25	32	82.1	136	2 D73805	cytochrome c552 -
26	32	82.1	166	2 B70369	transcription anti
27	32	82.1	178	2 A93235	transcription anti
28	32	82.1	183	2 B98099	hypothetical prote
29	32	82.1	411	2 H97652	

30	32	82.1	411	2 AF2876	conserved hypothet
31	32	82.1	472	2 S41720	intermediate filam
32	32	82.1	474	2 JC2037	6-phosphofructo-2-
33	32	82.1	475	2 T40840	pyruvate kinase (E
34	32	82.1	519	2 S74242	6-phosphofructo-2-
35	32	82.1	532	1 QPGL	neurofilament trip
36	32	82.1	628	2 A70376	conserved hypothet
37	32	82.1	693	2 T05006	hypothetical prote
38	32	82.1	708	2 S70834	transcription init
39	32	82.1	784	2 P40009	neurofilament trip
40	32	82.1	845	2 A45669	neurofilament trip
41	32	82.1	849	2 S00030	neurofilament trip
42	32	82.1	913	2 T52485	neurofilament prot
43	32	82.1	916	2 A27864	neurofilament trip
44	32	82.1	1110	2 I51116	NF-180 - sea lamp
45	32	82.1	1401	2 T48079	hypothetical prote

ALIGNMENTS

RESULT 1

QPM5L

neurofilament triplet L protein - mouse

N:Alternate names: 68k neurofilament protein; NF-L(low) protein; type IV IF protein

C:Species: Mus musculus (house mouse)

C>Date: 31-Mar-1988 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999

C:Accession: A25227; A26562; A43772; A41012; I55316

R:Lewis, S.A.; Cowan, N.J.

Mol. Cell. Biol. 6, 1529-1534, 1986

A:Title: Anomalous placement of introns in a member of the intermediate filament mult

A:Reference number: A25227; MUID:87064433; PMID:3785173

A:Accession: A25227

A:Molecule type: DNA

A:Residues: 1-543 <LEN>

A:Cross-references: GB:M13016; NID:9200023; PIDN:AAA39810.1; PID:9387492

A:Note: the authors translated the codon GGC for residue 5 as Ala, ACA for residue 88

1 as Glu

J. Cell Biol. 100, 843-850, 1985

A:Title: Genetics, evolution, and expression of the 68,000-mol-wt neurofilament prote

A:Reference number: A26562; MUID:8513134; PMID:3919033

A:Accession: A26562

A:Molecule type: mRNA

A:Residues: 242-543 <LEN>

A:Cross-references: GB:X02165

A:Experimental source: brain

R:Julien, J.P.; Meyer, D.; Flavell, D.; Hurst, J.; Grosfeld, F.

Brain Res. Mol. Brain Res. 1, 243-250, 1986

A:Title: Cloning and developmental expression of the murine neurofilament gene family

A:Reference number: A43772

A:Accession: A43772

A:Molecule type: mRNA

A:Residues: 1-5, 'Y', '7', '8', 'Y', '10-64', 'W', '66-72', 'L', '74-98', 'D', '100-194', 'R', '196-202, 204-239

A:Cross-references: GB:M20480; NID:9200037; PIDN:AAA39814.1; PID:9200038

A:Note: the authors translated the codon CGC for residue 195 as Ala

R:Shay, R.K.; Nixon, R.A.

J. Biol. Chem. 266, 18861-18867, 1991

A:Title: Identification of Ser-55 as a major protein kinase A phosphorylation site on

A:Reference number: A41012; MUID:92011653; PMID:1717455

A:Accession: A41012

A:Molecule type: protein

A:Residues: 52-57 <SIH>

R:Nakamura, K.; Ikenaka, K.; Maeda, K.; Tamura, T.

J. Biol. Chem. 265, 19786-19791, 1990

A:Title: Structure of the 68-kDa neurofilament gene and regulation of its expression.

A:Reference number: I55316; MUID:91060592; PMID:2246261

A:Accession: I55316

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-5, 'Y', '7', '8', 'Y', '10-28 <RES>

A:Cross-references: GB:M55423; NID:9200027; PIDN:AAA39812.1; PID:9554245

C:Comment: This is the most abundant of the three neurofilament proteins and, as the

C:Genetics:
 A:Introns: 349/3; 391/2; 498/1
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:2-72/Domain: head <HED>
 F:94-125/Domain: coil 1a, alpha-helical rod #status predicted <R1a>
 F:126-138/Region: linker 1
 F:139-234/Domain: coil 1b, alpha-helical rod #status predicted <R1b>
 F:235-256/Region: linker 12
 F:257-272/Domain: coil 2a, alpha-helical rod #status predicted <R2a>
 F:273-281/Region: linker 2
 F:282-401/Domain: coil 2b, alpha-helical rod #status predicted <R2b>
 F:404-543/Domain: tail <TAI>
 F:404-444/Region: tail subdomain a
 F:445-543/Region: tail subdomain b

Query Match 100.0%; Score 39; DB 1; Length 543;
 Best Local Similarity 100.0%; Pred. No. 9.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 Db 140 EOEIRDLR 147

RESULT 2
 S07144
 neurofilament triplet L protein - human
 N:Alternate names: neurofilament light polypeptide (66k)
 N:Contains: Glu-50 brain peptide
 C:Species: Homo sapiens (hmn)
 C:Date: 29-Jan-1993 #sequence, revision 29-Jan-1993 #text, change 21-Jul-2000
 C:Accession: S07144; I53832; A60703
 R:Julian, J.P.; Grosveld, F.; Yazdankhsh, K.; Flavell, D.; Meijer, D.; Mushynski, W.
 Biochem. Biophys. Acta 909, 10-20, 1987
 A:Title: The structure of a human neurofilament gene (NF-L): a unique exon-intron organ
 A:Reference number: S07144; MUID:87214213; PMID:3034332
 A:Accession: S07144

A:Molecule type: DNA
 A:Residues: 1-544 <JUL>
 A:Cross-references: EMBL:X05608; NID:q1495072; PIDN:CAA29697.1; PID:q1279504
 A:Note: the authors translated the codon ATG for residue 366 as Asn
 R:Pospelov, V.A.; Pospelova, T.V.; Julien, J.P.
 Cell Growth Differ. 5, 187-196, 1994
 A:Title: AP-1 and Krox-24 transcription factors activate the neurofilament light gene pr
 A:Reference number: I52832; MUID:94235564; PMID:8180132
 A:Accession: I52832

A:Molecule type: DNA
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Residues: 1-10 <POS>
 A:Cross-references: GB:S70309; NID:q547176; PIDN:AAD14057.1; PID:q4261757
 R:Nomata, Y.; Watanabe, T.; Wada, H.
 J. Biochem. 93, 825-831, 1993
 A:Title: Highly acidic proteins from human brain: purification and properties of Glu-50
 A:Reference number: A60703; MUID:83265667; PMID:6135695
 A:Accession: A60703

A:Molecule type: Protein
 A:Residues: 469-472, D, 474 <NDM>
 A:Experimental source: Glu-50 brain peptide
 A:Note: this acidic protein is named for its greater than fifty per cent glutamic acid c
 C:Genetics:
 A:Gene: GDB:NFL; NFL
 A:Cross-references: GDB:120227; OMIM:162280
 A:Map position: 8p21-8p21
 A:Introns: 349/3; 391/2; 498/1
 C:Superfamily: cytoskeletal keratin
 C:Keywords: brain; coiled coil; intermediate filament
 F:469-544/Product: Glu-50 peptide #status predicted <E50>

Query Match 100.0%; Score 39; DB 2; Length 544;
 Best Local Similarity 100.0%; Pred. No. 9.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 Db 138 EOEIRDLR 146

RESULT 3
 OEPGL
 neurofilament triplet L protein - pig
 N:Alternate names: 68k neurofilament protein
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 15-Nov-1984 #sequence, revision 28-May-1986 #text, change 10-Oct-1997
 C:Accession: A91337; A90973; A34569; A02963
 R:Geisler, N.; Plessmann, U.; Weber, K.
 FEBS Lett. 182, 475-478, 1985
 A:Title: The complete amino acid sequence of the major mammalian neurofilament prote
 A:Reference number: A91337; MUID:85154583; PMID:3920075
 A:Accession: A91337

A:Molecule type: protein
 A:Residues: 1-547 <GE1>
 R:Geisler, N.; Kaufmann, E.; Fischer, S.; Plessmann, U.; Weber, K.
 EMBO J. 2, 1295-1302, 1983
 A:Title: Neurofilament architecture combines structural principles of intermediate f
 A:Reference number: A90973

A:Accession: A90973
 A:Molecule type: protein
 A:Residues: 1-82; 278-548 <GE2>
 A:Note: residue 322 is either lysine or arginine
 R:Gonda, Y.; Nishizawa, K.; Ando, S.; Kitamura, S.; Minoura, Y.; Nish, Y.; Inagaki, I.
 Biochem. Biophys. Res. Commun. 167, 1316-1325, 1990
 A:Title: Involvement of protein kinase C in the regulation of assembly-disassembly c
 A:Reference number: A34569; MUID:90211318; PMID:2108674
 A:Accession: A34569

A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 9-14; 23-29; 30-53 <GON>
 A:Comment: Mammalian neurofilaments usually contain three polypeptides, L, M, and H
 ke all other intermediate filament proteins: a conserved alpha-helical region, whose
 C:Comment: The amino-terminal headpiece is basic with a high content of hydroxyamino
 C:Comment: The extra mass and high charge density that distinguish the neurofilament
 charged scaffolding structure suitable for interaction with other neuronal componen
 C:Comment: The boundaries of the domains between residues 70-92 and 399-402 are not
 C:Comment: this protein was isolated from spinal cord.
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:1-70/Domain: head <HED>
 F:92-123/Domain: coil 1a, alpha-helical rod <R1a>
 F:137-232/Domain: coil 1b, alpha-helical rod <R1b>
 F:255-399/Domain: coil 2, alpha-helical rod <R2>
 F:402-548/Domain: tail <TAI>
 F:402-442/Region: tail subdomain a
 F:443-548/Region: tail subdomain b

Query Match 100.0%; Score 39; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 Db 138 EOEIRDLR 145

RESULT 4
 JN0094
 neurofilament protein-L - bovine
 N:Alternate names: NF-L
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 03-Jul-1998 #sequence, revision 10-Jul-1998 #text, change 07-May-1999
 C:Accession: JN0094
 R:Hashimoto, R.; Nakamura, Y.; Goto, H.; Wada, Y.; Sakoda, S.; Kalibuchi, K.; Inagaki, I.
 Biochem. Biophys. Res. Commun. 245, 407-411, 1998
 A:Title: Domain- and site-specific phosphorylation of bovine NF-L by Rho-associated
 A:Reference number: JN0094; MUID:98238650; PMID:9571164

A:Accession: J00094
 A:Molecule type: protein
 A:Residues: 1-554 <HAS>
 C:Comment: Domain- and site-specific phosphorylation by Rho-kinase regulates the assembly
 C:Superfamily: cytoskeletal keratin

Query Match 100.0%; Score 39; DB 2; Length 554;
 Best Local Similarity 100.0%; Pred. No. 9.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 |||||
 DB 139 EOEIRDLR 146

RESULT 5
 A0785
 hypothetical protein, 24k - rice tungro bacilliform virus

N:Alternate names: hypothetical protein 1
 C:Species: rice tungro bacilliform virus, RTBV
 C:Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 11-Jan-2002

C:Accession: A40785; S16666
 R:Ou, R.; Bhattacharyya, M.; Laco, G.S.; de Kochko, A.; Rao, B.L.S.; Kanlewska, M.B.; El
 Virolgio 185, 354-364, 1991

A:Title: Characterization of the genome of rice tungro bacilliform virus: comparison with
 A:Reference number: A40785; MUID:92024093; PMID:1926781

A:Accession: A40785
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-221 <OUR>
 A:Cross-references: GB:M55026

A>Note: readthrough of the terminator TAA occurs between codons ACA for 22-Thr and GGT for
 R:Hay, J.M.; Jones, M.C.; Blakebrough, M.L.; Dasgupta, I.; Davies, J.W.; Hull, R.

Nucleic Acids Res. 19, 2615-2621, 1991
 A:Title: An analysis of the sequence of an infectious clone of rice tungro bacilliform virus

A:Reference number: S16666; MUID:91252204; PMID:2041739

A:Accession: S16666
 A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 23-116, 'F', 118-221 <HAY>

A:Cross-references: EMBL:X57924; NID:961910; PIDN:CAA40995.1; PID:91334939
 A:Experimental source: Philippine isolate

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1991

Query Match 92.3%; Score 36; DB 2; Length 221;
 Best Local Similarity 87.5%; Pred. No. 14;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 |||||
 DB 203 EOEIRDLR 210

RESULT 6
 S6546
 phosphoenolpyruvate carboxykinase (GTP) (EC 4.1.1.32) precursor, mitochondrial - human

C:Species: Homo sapiens (man)
 C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000

C:Accession: S6546
 R:Modaresi, S.; Christ, B.; Bratke, J.; Zahn, S.; Heise, T.; Jungermann, K.

Biochem. J. 315, 807-814, 1996
 A:Title: Molecular cloning, sequencing and expression of the cDNA of the mitochondrial

A:Reference number: S6546; MUID:96220091; PMID:8645161

A:Accession: S6546
 A:Status: preliminary
 A:Molecule type: mRNA

A:Residues: 1-640 <MOD>
 A:Cross-references: EMBL:X92720; NID:91403049; PIDN:CAA63380.1; PID:91403050

A>Note: The authors translated the codon GCC for residue 477 as Arg

C:Genetics: nuclear
 A:Genome: nuclear
 C:Superfamily: phosphoenolpyruvate carboxykinase (GTP)
 C:Keywords: carbon-carbon lyase; carboxy-lyase; mitochondrion

Query Match 92.3%; Score 36; DB 2; Length 640;
 Best Local Similarity 75.0%; Pred. No. 43;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 |||||
 DB 604 EOEIRDLR 611

RESULT 7
 VEMSGF
 glial fibrillary acidic protein, astrocyte - mouse

N:Alternate names: GFAP
 C:Species: Mus musculus (house mouse)
 C:Date: 13-Aug-1986 #sequence_revision 30-Jun-1993 #text_change 10-Dec-1999

C:Accession: B60052; A24343; A02960; S42464; I49688
 R:Brenner, M.; Lampel, K.; Nakatani, Y.; Mill, J.; Banner, C.; Mearrow, K.; Dohadwala,

Brain Res. Mol. Brain Res. 7, 277-286, 1990
 A:Title: Characterization of human cDNA and genomic clones for glial fibrillary acid

A:Reference number: B60052; MUID:90294716; PMID:2163003

A:Accession: B60052
 A:Status: not compared with conceptual translation

A:Molecule type: DNA
 A:Residues: 1-57

A>Note: the authors present evidence that the sequence in reference A24343 contains a
 R:Balcarek, J.M.; Cowan, N.J.

Nucleic Acids Res. 13, 5527-5543, 1985
 A:Title: Structure of the mouse glial fibrillary acidic protein gene: implications for

A:Reference number: A24343; MUID:85297756; PMID:2994002

A:Accession: A24343
 A:Molecule type: DNA

A:Residues: 'MPPRRMSAS', 23-346, 'H', 348-430 <BAL>
 A:Cross-references: GB:X02801; NID:951065; PIDN:CAA26571.1; PID:951066

A>Note: the authors translated the codon CAT for residue 174 as Asp, ACC for residue
 R:Lewis, S.A.; Balcarek, J.M.; Krek, V.; Shelanski, M.; Cowan, N.J.

Proc. Natl. Acad. Sci. U.S.A. 81, 2743-2746, 1984
 A:Title: Sequence of a cDNA clone encoding mouse glial fibrillary acidic protein: str

A:Reference number: A02960; MUID:84194082; PMID:6585825

A:Accession: A02960
 A:Molecule type: mRNA

A:Residues: 28-430 <LEW>
 A:Cross-references: GB:X01347; NID:9193465; PIDN:AAA37678.1; PID:9387163

A:Experimental source: clone G1
 A>Note: the authors translated the codon CAT for residue 147 as Asp

R:Ratton, J.E.; Lu, X.; Hutcheon, A.M.; Quinlan, R.A.
 submitted to the EMBL Data Library, March 1994

A:Description: Identification of two non-alpha-helical domain motifs important in the

A:Reference number: S42464

A:Accession: S42464
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-26 <RAI>

A:Cross-references: EMBL:X70141; NID:9460826; PIDN:CAA55020.1; PID:9460827
 R:Cowan, N.J.; Lewis, S.A.; Balcarek, J.M.; Krek, V.; Shelanski, M.L.

Ann. N. Y. Acad. Sci. 455, 575-582, 1985
 A:Title: Structural implications of a cDNA clone encoding mouse glial fibrillary acid

A:Reference number: I49688; MUID:86104618; PMID:3866511

A:Accession: I49688
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 28-430 <RES>

A:Cross-references: GB:M25937; NID:9193467; PIDN:AAA37679.1; PID:9387164
 C:Comment: GFAP is a cell-specific marker that, during the development of the central

C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:1-19/Domain: head (fragment) #status predicted <HED>
 F:20-348/Domain: rod #status predicted <ROD>
 F:349-403/Domain: tail #status predicted <TLE>

OY 1 EOEIRDLR 8
 ||:|||||
 Db 202 EEEVRDLR 209

RESULT 8

A44841
 low molecular weight neuronal intermediate filament protein XNIF - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
 C:Accession: A44841
 R:Charnas, L. R.; Szaro, B. G.; Gainer, H.
 J. Neurosci. 12, 3010-3024, 1992
 A:Title: Identification and developmental expression of a novel low molecular weight neuronal intermediate filament protein XNIF
 A:Reference number: A44841; MUID:92356194; PMID:1494944
 A:Accession: A44841
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-469 <CH>
 A:Experimental source: brain
 A:Note: sequence extracted from NCBI backbone (NCBIN:110223, NCBIPI:110224)
 C:Superfamily: cytoskeletal keratin

Query Match 89.7%; Score 35; DB 2; Length 469;
 Best Local Similarity 87.5%; Pred. No. 48;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 ||:|||||
 Db 137 EOEIRDLR 144

RESULT 9

KIHUP1
 pyruvate kinase (EC 2.7.1.40), hepatic splice form L - human
 N:Alternate names: PK-L-type isozyme; pyruvate kinase isoform L (liver)
 C:Species: Homo sapiens (man)
 C>Date: 01-Dec-1989 #sequence_revision 05-Dec-1997 #text_change 11-Jun-1999
 C:Accession: I52269; A30150; A29414; FN0451
 R:Kanno, H.; Fujii, H.; Miwa, S.
 Biochem. Biophys. Res. Commun. 188, 516-523, 1992
 A:Title: Structural analysis of human pyruvate kinase L-gene and identification of the R
 A:Reference number: I52269; MUID:93075125; PMID:1445295
 A:Accession: I52269
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL
 A:Molecule type: DNA
 A:Residues: 1-98, 'T', '100-200', 'S', '202-349', 'P', '351-391', 'R', '393-543' <TRAN1>
 A:Cross-references: EMBL:M15465; NID:9220048; PID:9220050
 A:Note: following the authors' interpretation the Genbank translation starts with the R
 A:Title: Human liver type pyruvate kinase: complete amino acid sequence and the expressi
 A:Reference number: A30150; MUID:88158079; PMID:3126495
 A:Accession: A30150
 A:Molecule type: mRNA
 A:Residues: 1-98, 'T', '100-200', 'S', '202-349', 'P', '351-391', 'R', '393-543' <TRAN1>
 A:Cross-references: EMBL:M15465; NID:9220048; PID:9220050
 R:Tanai, K.; Fujii, H.; Tsubumi, H.; Sukegawa, J.; Toyoshima, K.; Yoshida, M.C.; Noguchi
 Biochem. Biophys. Res. Commun. 143, 431-438, 1987
 A:Title: Human liver type pyruvate kinase: cDNA cloning and chromosomal assignment.
 A:Reference number: A29414; MUID:87184521; PMID:3566732
 A:Accession: A29414
 A:Molecule type: mRNA
 A:Residues: 439-543 <TRAN2>
 A:Cross-references: GB:J03640; NID:9189995
 C:Genetics:
 A:Gene: GDB:PKLR
 A:Cross-references: GDB:120294; OMIM:266200
 A:Map position: 1q21-1q21
 A:Introns: 3/1; 64/1; 94/3; 138/3; 201/1; 291/2; 341/3; 392/3; 448/2; 509/1
 A:Note: alternative promoters are used for the erythrocyte (R) and hepatic (L) forms
 C:Complex: homotetramer

C:Function:
 A:Description: catalyzes the transphosphorylation of phosphoenolpyruvate and ADP to
 A:Pathway: glycolysis; gluconeogenesis
 A:Superfamily: pyruvate kinase

C:Keywords: alternative initiators; alternative splicing; ATP biosynthesis; glucone
 F:12/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status pre
 F:85,255,306/Binding site: substrate phosphate (Arg, Ser, Arg) #status predicted
 F:282/Active site: Lys #status predicted
 F:284,305,306/Binding site: magnesium (Glu, Ala, Arg) #status predicted
 F:341,376/Binding site: potassium (Glu, Glu) #status predicted

Query Match 89.7%; Score 35; DB 1; Length 543;
 Best Local Similarity 75.0%; Pred. No. 56;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 ||:|||||
 Db 235 EDDVRDLR 242

RESULT 10

B46024
 neurofilament-L subunit - quail
 C:Species: Coturnix coturnix (quail)
 C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
 C:Accession: B46024
 R:Ohara, O.; Gahara, Y.; Miyake, T.; Teraoka, H.; Kitamura, T.
 J. Cell Biol. 121, 387-395, 1993
 A:Title: Neurofilament deficiency in quail caused by nonsense mutation in neurofilam
 A:Reference number: A46024; MUID:93224534; PMID:8468353
 A:Accession: B46024
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-556 <OH>
 A:Experimental source: subsp. japonica, TRP
 A:Note: sequence inconsistent with the nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:129455, NCBIPI:129482)
 C:Superfamily: cytoskeletal keratin

Query Match 89.7%; Score 35; DB 2; Length 556;
 Best Local Similarity 87.5%; Pred. No. 57;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 ||:|||||
 Db 141 EOEIRDLR 148

RESULT 11

KIHUPR
 pyruvate kinase (EC 2.7.1.40), erythrocyte splice form R - human
 N:Alternate names: PK-R-type isozyme; pyruvate kinase isoform R (erythroid)
 C:Species: Homo sapiens (man)
 C>Date: 01-Dec-1989 #sequence_revision 05-Dec-1997 #text_change 26-Feb-1999
 C:Accession: I52269; A40991; FN0451
 R:Kanno, H.; Fujii, H.; Miwa, S.
 Biochem. Biophys. Res. Commun. 188, 516-523, 1992
 A:Title: Structural analysis of human pyruvate kinase L-gene and identification of t
 A:Reference number: I52269; MUID:93075125; PMID:1445295
 A:Accession: I52269
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB
 A:Molecule type: DNA
 A:Residues: 1-93, 'VPITQCGADPQGRPREVCSGE', '34-574' <KAN>
 A:Cross-references: GB:DI3243; NID:9220048; PID:9220050
 A:Note: following the authors' interpretation the Genbank sequence uses an intron w1
 R:Kanno, H.; Fujii, H.; Hirono, A.; Miwa, S.
 Proc. Natl. Acad. Sci. U.S.A. 88, 8218-8221, 1991
 A:Title: cDNA cloning of human R-type pyruvate kinase and identification of a single
 A:Reference number: A40991; MUID:91376115; PMID:1896471
 A:Accession: A40991
 A:Molecule type: mRNA
 A:Residues: 1-37,570-574 <KA3>

A:Cross-references: GB:D90465
 A:Experimental source: reticulocyte cDNA library
 R:Kanno, H.; Fujii, H.; Tsujino, G.; Miwa, S.
 Biochem. Biophys. Res. Commun. 192, 46-52, 1993
 A:Title: Molecular basis of impaired pyruvate kinase isozyme conversion in erythroid cell
 A:Reference number: P00451, MUID:93235593, PMID:8476433
 A:Accession: P00451
 A:Molecule type: mRNA
 A:Residues: 365-380, 'P', 382-431 <KA2>
 C:Genetics:
 A:Gene: GDB:PKLR
 A:Cross-references: GDB:120294; OMIM:266200
 A:Map position: 1q21-q22
 A:Introns: 34/1; 95/1; 125/3; 169/3; 232/1; 322/2; 372/3; 423/3; 479/2; 540/1
 A:Note: alternative promoters are used for the erythrocyte (R) and hepatic (L) forms
 C:Complex: homotetramer
 C:Function:
 A:Description: catalyzes the transphosphorylation of phosphoenolpyruvate and ADP to pyruvate
 A:Pathway: glycolysis
 C:Superfamily: pyruvate kinase
 C:Keywords: alternative splicing; ATP biosynthesis; erythrocyte;
 F:43/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted
 F:116/286/337/Binding site: substrate phosphate (Arg, Ser, Arg) #status predicted
 F:313/Active site: lys #status predicted
 F:315/336/337/Binding site: magnesium (Glu, Ala, Arg) #status predicted
 F:372/407/Binding site: potassium (Glu, Glu) #status predicted

Query Match 89.7%; Score 35; DB 1; Length 574;
 Best Local Similarity 75.0%; Pred. No. 59;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDEIRDLR 8
 DB 266 EDDVDRLR 273

RESULT 12
 H83875
 two-component response regulator BH1808 [Imported] - Bacillus halodurans (strain C-125)
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: H83875
 R:Takami, H.; Nakase, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512362; PMID:11056132
 A:Accession: H83875
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-231 <STO>
 A:Cross-references: GB:AP001513; GB:BA000004; NID:G10174345; PIDN:BA05527.1; GSPDB:GNOC
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH1808
 C:Superfamily: ompr protein; response regulator homology

Query Match 87.2%; Score 34; DB 2; Length 231;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EDEIRDL 7
 DB 12 EDEIRDL 18

RESULT 13
 T51195
 hypothetical protein B7N4.20 [Imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 28-Jul-2000
 C:Accession: T51195
 R:Schulte, U.; Allyn, V.; Hohenseil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, July 2000

A:Reference number: Z25286
 A:Accession: T51195
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-497 <SCH>
 A:Cross-references: EMBL:AL390218; GSPDB:GN00116; NCSP:B7N4.20
 A:Experimental source: BAC clone B7N4; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B7N4.20
 A:Map position: 6
 A:Introns: 342/1

Query Match 87.2%; Score 34; DB 2; Length 497;
 Best Local Similarity 87.5%; Pred. No. 79;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EDEIRDLR 8
 DB 209 EDEIRDLR 216

RESULT 14
 B38932
 phospholipase C (PC 3.1.4.3) beta-1 form B - bovine (fragment)
 N:Alternate names: phospholipidase-specific phospholipase C isozyme C1; phospholipase
 C:Species: Bos primigenius laurus (cattle)
 C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 22-Jun-1998
 C:Accession: B38932; A38932; C38932; D38932; A48193; A45975; B45975; B48193; C45975
 R:Ferreira, P.A.; Shortridge, R.D.; Pak, W.L.
 submitted to GenBank, August 1993
 A:Reference number: A38932
 A:Accession: B38932
 A:Molecule type: mRNA
 A:Residues: 1-1023 <FER>
 A:Cross-references: GB:U3936; NID:G304240; PIDN:AAA30699.1; PID:G304241
 A:Experimental source: retina
 A:Accession: A38932
 A:Molecule type: mRNA
 A:Residues: 1-373, 386-1023 <FE2>
 A:Cross-references: GB:U3935; NID:G304238; PIDN:MAC37304.1; PID:G304239
 A:Experimental source: retina
 A:Accession: C38932
 A:Molecule type: mRNA
 A:Residues: 105-373, 386-1023 <FE3>
 A:Cross-references: GB:U3937; NID:G304242; PIDN:AAA30700.1; PID:G304243
 A:Experimental source: retina
 A:Accession: D38932
 A:Molecule type: mRNA
 A:Residues: 105-1023 <FE4>
 A:Cross-references: GB:U3938; NID:G304244; PIDN:AAA30701.1; PID:G304245
 A:Experimental source: retina
 R:Ferreira, P.A.; Shortridge, R.D.; Pak, W.L.
 Proc. Natl. Acad. Sci. U.S.A. 90, 6042-6046, 1993
 A:Title: Distinctive subtypes of bovine phospholipase C that have preferential expres
 A:Reference number: A48193; MUID:933117620; PMID:8327481
 A:Accession: A48193
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-300, 412-670 <FE5>
 A:Cross-references: GB:U3935
 R:Min, D.S.; Kim, D.M.; Lee, Y.H.; Seo, J.; Suh, P.G.; Ryu, S.H.
 J. Biol. Chem. 268, 12207-12212, 1993
 A:Title: Purification of a novel phospholipase C isozyme from bovine cerebellum.
 A:Reference number: A45975; MUID:93280199; PMID:8389372
 A:Accession: A45975
 A:Molecule type: protein
 A:Residues: 20-23, 'D', 25-28, 'X', 30-31, 'DX', 34-35, 'L', 37, 259, 'X', 261-269, 315-327, 'I', 3
 A:Experimental source: cerebellum
 A:Note: sequence modified after extraction from NCBI backbone
 C:Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phospholipase I; 1-phosphat
 odiesterase domain y homology
 C:Keywords: alternative splicing; phosphoric diester hydrolase
 F:1-1023/Product: phospholipase C beta-1 form B (fragment) #status predicted <MAT>

F:1-373,386-1023/Product: phospholipase C beta-1 form A (fragment) #status predicted <NA>
 F:105-1023/Product: phospholipase C beta-II form B #status predicted <NA4>
 F:105-373,386-1023/Product: phospholipase C beta-II form A #status predicted <NA3>
 F:151-299/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X hom
 F:412-532/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain Y hom

Query Match 87.2% Score 34; DB 2; Length 1023;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDL 7
 |||||
 DB 883 EOEIRDL 889

RESULT 15

A48047
 phospholipase C (EC 3.1.4.-) beta-4 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 02-Jun-2000
 C:Accession: A48047; JN0691
 R:Lee, C.W.; Park, D.J.; Lee, K.H.; Kim, C.G.; Rhee, S.G.
 J. Biol. Chem. 268, 21318-21327, 1993
 A:Title: Purification, molecular cloning, and sequencing of phospholipase C-beta 4.
 A:Reference number: A48047; MUID:94012687; PMID:8407970
 A:Accession: A48047
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-1176 <LEF>
 A:Cross-references: GB:L15556; NID:9404071
 A:Experimental source: brain
 A:Note: sequence extracted from NCBI backbone (NCBIP:138517)
 A:Note: this translation is not annotated in GenBank entry RABETA4AA, release 113.0
 R:Kim, M.J.; Bahk, Y.Y.; Min, D.S.; Lee, S.J.; Ryu, S.H.; Suh, P.G.
 Biochem. Biophys. Res. Commun. 194, 706-712, 1993
 A:Title: Cloning of cDNA encoding rat phospholipase C-beta4, a new member of the phospho
 A:Reference number: JN0691; MUID:9343926; PMID:7688223
 A:Accession: JN0691
 A:Molecule type: mRNA
 A:Residues: 1-254; 'M', 256-307, 'A', 309-416, 'E', 418-469, 'K', 471-504, 506-545, 'DE', 548-734, '
 'M', 1045-1176 <KIM>
 A:Cross-references: GB:L18962
 A:Experimental source: brain
 A:Note: the authors translated the codon AAC for residue 140 as Lys, CAG for residue 102
 C:Comment: This enzyme belongs a member of the phosphoinositide specific phospholipase C
 messenger molecules. Inositol 1,4,5-trisphosphate and diacylglycerol in response to the
 C:Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase I; 1-phosphatid
 odiesterase domain Y homology
 C:Keywords: phosphoric diester hydrolase
 F:315-463/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X hom
 F:565-685/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain Y hom

Query Match 87.2% Score 34; DB 1; Length 1176;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDL 7
 |||||
 DB 1036 EOEIRDL 1042

Search completed: November 13, 2002, 13:23:31
 Job time : 13.5684 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 6.6383 Seconds

(without alignments)
49,984 Million cell updates/sec

Title: US-09-856-086-7

Perfect score: 39

Sequence: 1 EOEIRDLR 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476338 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	541	1 NFL_RAT	P19527 rat mus nov
2	39	100.0	542	1 NFL_MOUSE	P08551 mus musculi
3	39	100.0	543	1 NFL_HUMAN	P07196 homo sapien
4	39	100.0	548	1 NFL_PIG	P02547 sus scrofa
5	39	100.0	554	1 NFL_BOVIN	P02548 bos taurus
6	36	92.3	199	1 YP24_RTBVP	P27300 rice tungro
7	36	92.3	640	1 PCOM_HUMAN	P16822 homo sapien
8	35	89.7	430	1 GPAP_MOUSE	P03995 mus musculi
9	35	89.7	470	1 XNIF_XENLA	P35617 xenopus lae
10	35	89.7	555	1 NFL_CONJA	P00216 coturnix co
11	35	89.7	574	1 KPIR_HUMAN	P30613 homo sapien
12	34	87.2	1023	1 PIR4_BOVIN	P07722 bos taurus
13	34	87.2	1175	1 PIR4_HUMAN	P01517 homo sapien
14	34	87.2	1175	1 PIR4_RAT	P09407 rattus norv
15	33	84.6	458	1 IEF3_TORCA	P23729 torpedo cal
16	33	84.6	464	1 IFE_BRALA	P00448 brachyosteo
17	33	84.6	525	1 ASF2_YEAST	P32448 saccharomyc
18	33	84.6	657	1 LAMA_CHICK	P13548 gallus gall
19	33	84.6	857	1 NFM_CHICK	P16053 gallus gall
20	32	82.1	136	1 VG46_BPMD2	P06437 mycobacteri
21	32	82.1	454	1 NFM_PIG	P08552 sus scrofa
22	32	82.1	475	1 KPYK_CORGL	P04678 corynebacte
23	32	82.1	519	1 P262_MOUSE	P07065 mus musculi
24	32	82.1	557	1 F262_RAT	P09115 rattus norv
25	32	82.1	628	1 DKS_AQAAE	P07036 aquifex aeo
26	32	82.1	708	1 RPSD_MYXA	P17531 myxococcus
27	32	82.1	790	1 KIP9_MOUSE	P09404 mus musculi
28	32	82.1	810	1 NFM_BOVIN	P12339 rattus norv
29	32	82.1	845	1 NFM_RAT	P08553 mus musculi
30	32	82.1	848	1 NFM_MOUSE	P07197 homo sapien
31	32	82.1	915	1 NFM_HUMAN	P23133 pseudomonas
32	31	79.5	276	1 TODF_PSEPU	P32153 escherichia
33	31	79.5	356	1 FRVX_ECOLI	

34	31	79.5	414	1 CDV1_MOUSE	O35594 mus musculi
35	31	79.5	426	1 YK96_AERPE	O94a42 aeropyrum p
36	31	79.5	461	1 VIME_ONCMY	P48674 oncorhynchus p
37	31	79.5	469	1 TRPC_PASMU	P57855 pasturella
38	31	79.5	544	1 NFL_XENLA	P35616 xenopus lae
39	31	79.5	650	1 YB73_MYCTU	Q10533 mycobacteri
40	31	79.5	716	1 REP_DROME	P41073 drosophila
41	31	79.5	1024	1 CARB_DEIRA	O94w40 deinococcus
42	31	79.5	1509	1 MYSN_ACACA	P05659 acanthamoeb
43	31	79.5	2035	1 EYPL_MOUSE	O94952 mus musculi
44	30	76.9	184	1 KY17_YEAST	O02205 saccharomyc
45	30	76.9	228	1 VP25_NPVAC	P41483 autographa

ALIGNMENTS

RESULT 1

NFL_RAT ID NFL_RAT STANDARD: PRT: 541 AA.

AC P19527: 063367; 01-FEB-1991 (Rel. 17, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L).
 GN NEFL OR NFL OR NF68
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP MEDLINE=90184052; PubMed=2516804;
 RX Chin S.S., Liem R.K.H., "Expression of rat neurofilament proteins NF-L and NF-M in
 RT transfection non-neuronal cells.";
 RL Eur. J. Cell Biol. 50:475-490(1989).
 RN [2]
 RP SEQUENCE OF 197-483 FROM N.A.
 RX MEDLINE=85252830; PubMed=3925999;
 RA Jullen J.-P., Ramchandran K., Grosvelde F.,
 RT "Cloning of a cDNA encoding the smallest neurofilament protein from
 RL the rat.";
 RN Biochim. Biophys. Acta 825:398-404(1985).
 RN [3]
 RP SEQUENCE OF 1-10 FROM N.A.
 RX MEDLINE=95264348; PubMed=7745611;
 RA Reeben M., Neuman T., Palgi J., Palm K., Paalme V., Saarna M.,
 RT "Characterization of the rat light neurofilament (NF-L) gene promoter
 RL and identification of NGF and cAMP responsive regions.";
 RN J. Neurosci. Res. 40:177-188(1995).
 RN [4]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=93346421; PubMed=8344946;
 RA Dong D.L.-Y., Xu Z.-S., Chevalier M.R., Cotter R.J., Cleveland D.W.,
 RT Hart G.W.,
 RL "Glycosylation of mammalian neurofilaments. Localization of multiple
 RT O-linked N-acetylglucosamine moieties on neurofilament polypeptides
 RL L and M.";
 RN J. Biol. Chem. 268:16679-16687(1993).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- PTM: O-GLYCOSYLATED; CONTAINS THREE N-ACETYLGLUCOSAMINE SIDE
 CC CHAINS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM

FILAMENTS.

-1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: AF031880; AAB87069.1; -
 DR EMBL: M25638; AAA1694.1; -
 DR EMBL: X53981; CAA37931.1; -
 DR PIR: A21762; A21762.
 DR GlycoSuiteDB: P19527; -
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 DR Intermediate filament; Colled coll; Neurone; Glycoprotein.
 KW Intermediate filament; Colled coll; Neurone; Glycoprotein.
 FT INIT_MET 0
 FT DOMAIN 1 92
 FT DOMAIN 93 396
 FT DOMAIN 397 541
 FT DOMAIN 541 93
 FT DOMAIN 125 137
 FT DOMAIN 138 233
 FT DOMAIN 234 252
 FT DOMAIN 253 280
 FT DOMAIN 272 280
 FT DOMAIN 281 396
 FT DOMAIN 397 443
 FT DOMAIN 444 541
 FT CARBOHYD 20
 FT CARBOHYD 20
 FT CARBOHYD 26
 FT SITE 381 391
 FT FT
 FT CONFLICT 197 202
 FT CONFLICT 399 399
 FT CONFLICT 476 476
 FT CONFLICT 480 483
 FT SEQUENCE 541 AA; 61204 MW; 0D17839AF226918A CRC64;
 Query Match 100.0%; Score 39; DB 1; Length 541;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EOEIRDLR 8
 Db 139 EOEIRDLR 146
 RESULT 2
 NFI_MOUSE STANDARD: PRT: 542 AA.
 ID NFI_MOUSE
 AC P08531;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L).
 GN NFI OR NFI.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=87064433; PubMed=3785173;
 RA Lewis S.A., Cowan N.J.;

"Anomalous placement of introns in a member of the intermediate filament multigene family: an evolutionary conundrum";
 Mol. Cell. Biol. 6:1529-1534(1986).

[2] SEQUENCE FROM N.A.

RC TISSUE-Brain;
 RX MEDLINE=87156637; PubMed=3103856;
 RA Julien J.-P., Meyer D., Flavell D., Hurst J., Grosved F.;
 RT "Cloning and developmental expression of the murine neurofilament gene family";
 RL Brain Res. 387:243-250(1986).

[3] SEQUENCE OF 241-542 FROM N.A.

RC TISSUE-Brain;
 RX MEDLINE=8513134; PubMed=3919033;
 RA Lewis S.A., Cowan N.J.;
 RT "Genetics, evolution, and expression of the 68,000-mol-wt neurofilament protein: isolation of a cloned cDNA probe";
 RL J. Cell Biol. 100:843-850(1985).

[4] SEQUENCE OF 1-27 FROM N.A.

RP MEDLINE=91060592; PubMed=2246261;
 RX Nakahira K., Ikenaka K., Wada K., Tamura T.A., Furutachi T.,
 RA Mikoshiba K.;
 RT "Structure of the 68-kDa neurofilament gene and regulation of its expression";
 RL J. Biol. Chem. 265:19786-19791(1990).

-1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M, AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.

-1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM A CHARGED SCAPFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH OTHER NEURONAL COMPONENTS OR IONS.
 -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPYTHETIAL INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM FILAMENTS.

-1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X02165; CAB51616.1; -
 DR EMBL: M20480; AAA39814.1; -
 DR EMBL: M13016; AAA39810.1; -
 DR EMBL: M55423; AAA39812.1; -
 DR PIR: A25227; QFMSL.
 DR MGD: MGI:97313; NFI.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 DR Intermediate filament; Colled coll; Neurone; Glycoprotein.
 KW Intermediate filament; Colled coll; Neurone; Glycoprotein.
 FT INIT_MET 0
 FT DOMAIN 1 92
 FT DOMAIN 93 396
 FT DOMAIN 397 542
 FT DOMAIN 542 93
 FT DOMAIN 125 137
 FT DOMAIN 138 233
 FT DOMAIN 234 252
 FT DOMAIN 253 271
 FT DOMAIN 272 280
 FT DOMAIN 281 396
 FT DOMAIN 397 443
 FT DOMAIN 444 542
 FT CARBOHYD 20
 FT CARBOHYD 26
 FT CARBOHYD 26
 FT O-LINKED (GLCNAC) (BY SIMILARITY).
 O-LINKED (GLCNAC) (BY SIMILARITY).

```

FT SITE 381 391 EPIPOPE (RECOGNIZED BY IF-SPECIFIC
FT CONFLICT 5 5 MONOCLONAL ANTIBODY).
FT CONFLICT 8 8 Y -> S (IN REF. 1).
FT CONFLICT 8 8 Y -> I (IN REF. 1).
FT CONFLICT 64 64 M -> K (IN REF. 1).
FT CONFLICT 72 72 V -> L (IN REF. 2).
FT CONFLICT 98 98 D -> H (IN REF. 1).
FT CONFLICT 194 194 R -> A (IN REF. 1).
FT CONFLICT 202 202 MISSING (IN REF. 2).
FT CONFLICT 239 239 Y -> I (IN REF. 1).
SQ SEQUENCE 542 AA; 61448 MW; 8EE9B8C6F0831D8C CRC64;

Query Match 100.0%; Score 39; DB 1; Length 542;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 139 EOEIRDLR 146

RESULT 3
NFL_HUMAN STANDARD; PRT; 543 AA.
ID P07196; 016154;
AC 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
GN NEFL OR NEFL OR NF68.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87214213; PubMed=3034332;
RA Julien J.-P., Grosfeld F., Yazdankhah K., Flavell D., Meljer D.,
RA Mushynski W.;
RT "The structure of a human neurofilament gene (NF-L): a unique exon-
RT intron organization in the intermediate filament gene family.";
RL Blohm. Biophys. Acta 909:10-20(1987).
RN [2]
RP SEQUENCE OF 1-10 FROM N.A.
RX MEDLINE=9423564; PubMed=8180132;
RA Pospelov V.A., Pospelova T.V., Julien J.-P.;
RT "Ap-1 and Krox-24 transcription factors activate the neurofilament
RT light gene promoter in p19 embryonal carcinoma cells.";
RL Cell Growth Differ. 5:187-196(1994).
RN [3]
RP VARIANT CMT2E PRO-332.
RX MEDLINE=20307176; PubMed=10841809;
RA Merilanova I.V., Perepelov A.V., Polyakov A.V., Sitnikov V.F.,
RA Dadali E.L., Oparin R.B., Petlin A.N., Evgrafov O.V.;
RT "A new variant of Charcot-Marie-Tooth disease type 2 is probably the
RT result of a mutation in the neurofilament-light gene.";
RL Am. J. Hum. Genet. 67:37-46(2000).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- DISEASE: DEFECTS IN NEFL ARE A CAUSE OF CHARCOT-MARIE-TOOTH
CC DISEASE TYPE 2E (CMT2E).
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC INTERMEDIATE FILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -1- DATABASE: NAME-INHERITED peripheral neuropathies mutation db;

```

```

CC WWW="http://molgen-www.uia.ac.be/CMTmutations/".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
CC EMBL; X05608; CAA29097.1; -
CC EMBL; S70309; AAD14057.1; -
CC PIR; S07144; S07144.
CC Genew; HGNC:7739; NEFL.
CC MIM; 162280; -
CC DR InterPro; IPR001664; IF.
CC DR Pfam; PF00036; Filament; 1.
CC DR PROSITE; PS00226; IF; 1.
CC KW Intermediate filament; Coiled coil; Neurone; Glycoprotein;
CC Disease mutation; Charcot-Marie-Tooth disease.
CC FT INT_MET 0 0
CC FT DOMAIN 1 91 HEAD.
CC FT DOMAIN 92 396 ROD.
CC FT DOMAIN 397 543 TAIL.
CC FT DOMAIN 92 123 COIL 1A.
CC FT DOMAIN 124 136 LINKER 1.
CC FT DOMAIN 137 234 COIL 1B.
CC FT DOMAIN 235 252 LINKER 12.
CC FT DOMAIN 253 271 COIL 2A.
CC FT DOMAIN 272 280 LINKER 2.
CC FT DOMAIN 281 396 COIL 2B.
CC FT DOMAIN 397 443 TAIL, SUBDOMAIN B (ACIDIC).
CC FT DOMAIN 444 543 TAIL, SUBDOMAIN A (ACIDIC).
CC FT CARBOHD 20 20 O-LINKED (GLCNAc) (BY SIMILARITY).
CC FT CARBOHD 26 26 O-LINKED (GLCNAc) (BY SIMILARITY).
CC FT SITE 381 391 EPIPOPE (RECOGNIZED BY IF-SPECIFIC
CC FT VARIANT 332 332 Q -> P (IN CMT2E).
CC FT SEQUENCE 543 AA; 61645 MW; 7A0F1ADB5BED22F6 CRC64;
CC SO

Query Match 100.0%; Score 39; DB 1; Length 543;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EOEIRDLR 8
DB 138 EOEIRDLR 145

RESULT 4
NFL_PIG STANDARD; PRT; 548 AA.
ID P02547;
AC P02547;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
GN NEFL.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=85154583; PubMed=3920075;
RA Geisler N., Plessmann U., Weber K.;
RT "The complete amino acid sequence of the major mammalian
RT neurofilament protein (NF-L).";
RL FEBS Lett. 182:475-478(1985).
RN [2]
RP SEQUENCE OF 1-82 AND 278-548.

```

RA Geisler N., Kaufmann E., Fischer S., Plessmann U., Weber K.;
 RT "Neurofilament architecture combines structural principles of
 RT intermediate filaments with carboxy-terminal extensions increasing
 RT in size between triplet proteins.";
 RL EMBL 3:1295-1302(1983).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 DR PIR: A02963; OPRCL.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; Filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neurone; Glycoprotein.
 FT DOMAIN 1 91 HEAD.
 FT DOMAIN 92 395 ROD.
 FT DOMAIN 396 548 TAIL.
 FT DOMAIN 92 123 COIL 1A.
 FT DOMAIN 124 136 LINKER 1.
 FT DOMAIN 137 232 COIL 1B.
 FT DOMAIN 233 251 LINKER 12.
 FT DOMAIN 252 270 COIL 2A.
 FT DOMAIN 271 279 LINKER 2.
 FT DOMAIN 280 395 COIL 2B.
 FT DOMAIN 396 442 TAIL, SUBDOMAIN A (ACIDIC).
 FT DOMAIN 443 548 TAIL, SUBDOMAIN B (ACIDIC).
 FT DOMAIN 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
 FT CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
 FT SITE 380 390 EPITOPE (RECOGNIZED BY IF-SPECIFIC
 FT MONOCLONAL ANTIBODY).
 FT UNSURE 322 322 OR K.
 SQ SEQUENCE 548 AA; 61940 MW; 83044813637AC739 CRC64;
 QY 1 EOETRLDR 8
 DB 138 EOETRLDR 145
 Query Match 100.0%; Score 39; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
 ID NFL_BOVIN STANDARD; PRT; 554 AA.
 AC P02548; P79127;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L) (Micro glutamic acid-rich
 DE protein).
 DE NEFL.
 GN NEFL.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Holstein; TISSUE-Brain;
 RA Hill W.D., Zhang L., Balin B.J., Sprinkle T.J.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 468-554.

RX MEDLINE-85154567; PubMed-3884373;
 RA Isebe T., Okuyama T.;
 RT "Brain micro glutamic acid-rich protein is the C-terminal endpiece of
 RT the neurofilament 68-kDa protein as determined by the primary
 RT sequence.";
 RL FEBS Lett. 182:389-392(1985).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 DR PIR: A02964; OPRCL.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; Filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neurone.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT DOMAIN 1 92 HEAD (BY SIMILARITY).
 FT DOMAIN 93 396 ROD (BY SIMILARITY).
 FT DOMAIN 397 554 TAIL (BY SIMILARITY).
 FT DOMAIN 92 124 COIL 1A.
 FT DOMAIN 125 137 LINKER 1.
 FT DOMAIN 138 233 COIL 1B.
 FT DOMAIN 234 252 LINKER 12.
 FT DOMAIN 253 271 COIL 2A.
 FT DOMAIN 272 280 LINKER 2.
 FT DOMAIN 281 396 COIL 2B.
 FT DOMAIN 397 443 TAIL, SUBDOMAIN A (ACIDIC).
 FT DOMAIN 444 554 TAIL, SUBDOMAIN B (ACIDIC).
 FT CONFLICT 494 500 MISSING (IN REF. 2).
 FT CONFLICT 509 509 A -> AEA (IN REF. 2).
 SQ SEQUENCE 554 AA; 62514 MW; D772B81CA2C3C1C1A CRC64;
 QY 1 EOETRLDR 8
 DB 139 EOETRLDR 146
 Query Match 100.0%; Score 39; DB 1; Length 554;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
 ID YP24_RTBVP STANDARD; PRT; 199 AA.
 AC P27500; P27530;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Hypothetical 24 kDa protein (P24) (ORF 1).
 DE Rice tungro bacilliform virus (isolate Philippines) (RTBV).
 OS Rice tungro bacilliform virus; Caulimoviridae;
 OC Viruses; Retroviruses; Caulimoviridae;
 OC Rice tungro bacilliform-like viruses.
 OC NCBI_TaxID=10655;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE-91252204; PubMed-2041739;
 RA Hay J.M., Jones M.C., Blakebrough M.L., Dasgupta I., Davies J.W.,
 RA Hull R.;
 RT "An analysis of the sequence of an infectious clone of rice tungro
 RT bacilliform virus, a plant pararetrovirus.";
 RT Nucleic Acids Res. 19:2615-2621(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92024093; PubMed-1926781;
 RA Ou R.D., Bhattacharyya M., Laco G.S., de Kochko A., Rao B.L.S.,
 RA Kanewski M.B., Elmer J.S., Rochester D.E., Smith C.E.,
 RA Beachy R.N.;
 RT Characterization of the genome of rice tungro bacilliform virus:
 RT comparison with Commelina yellow mottle virus and caulimoviruses";
 RT Virology 185:354-364(1991).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X57924; CAA40995.1; -
 CC DR EMBL: M65026; -; NOT_ANNOTATED_CDS.
 CC DR PIR: S16666; S16666.
 CC KM Hypothetical protein.
 CC FT CONFLICT 1 1 V -> F (IN REF. 2).
 CC FT CONFLICT 95 95 F -> S (IN REF. 2).
 CC FT SEQUENCE 199 AA; 23692 MW; 14B38CC64E80746 CRC64;
 SQ
 Query Match 92.3%; Score 36; DB 1; Length 199;
 Best Local Similarity 87.5%; Pred. No. 4.6;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EOEIRDLR 8
 DB 181 EOEIRDLR 188
 RESULT 7
 PRCR_HUMAN
 ID PRCR_HUMAN STANDARD; PRT; 640 AA.
 AC Q16822; Q43253; Q9BV62;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphoenolpyruvate carboxylase, mitochondrial precursor [GTP]
 DE (EC 4.1.1.32) (Phosphoenolpyruvate carboxylase) (PEPCK-M).
 GN PCK2 OR PEPCK2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE-96220691; PubMed-8645161;
 RA Modaresi S., Christ B., Bratke J., Zahn S., Heise T., Jungermann K.;
 RT "Molecular cloning, sequencing and expression of the cDNA of the
 RT mitochondrial form of phosphoenolpyruvate carboxylase from human
 RT liver.";
 RT Biochem. J. 315:807-814(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-98324897; PubMed-9651976;
 RA Modaresi S., Brechtel K., Christ B., Jungermann K.;
 RT "Human mitochondrial phosphoenolpyruvate carboxylase 2 gene.
 RT Structure, chromosomal localization and tissue-specific expression.";
 RT Biochem. J. 333:359-366(1998).
 RN [3]
 RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;
 RA Strausberg R.;
 RL Submitted (Dec-2000) to the EMBL/GenBank/DDBJ databases.
 CC -1- CATALYTIC ACTIVITY: GTP + oxaloacetate -> GDP + phosphoenolpyruvate
 CC + CO(2).
 CC -1- COFACTOR: REQUIRES MANGANESE (BY SIMILARITY).
 CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- DISEASE: DEFECTS IN PCK2 ARE A CAUSE OF HYPOGLYCEMIA.
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
 CC ONE AND A MITOCHONDRIAL ONE.
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE (GTP)
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X92720; CAA63380.1; -
 CC DR EMBL: Y11484; CAA72272.1; -
 CC DR EMBL: BC001454; AA01454.1; -
 CC DR SWISS-2DPAGE: Q16822; HUMAN.
 CC DR GeneW: HGNC:8725; PCK2.
 CC DR MIM: 261650; -
 CC DR InterPro: IPR000364; PEP_carboxykin.
 CC DR Pfam: PF00821; PEPCK; 1.
 CC DR ProDom: PD004738; PEP_carboxykin; 1.
 CC DR PROSITE: PS00505; PEPCK_GTP; 1.
 CC KM Gluconeogenesis; Lyase; Decarboxylase: GTP-binding; Mitochondrion;
 CC Transf. 1 32
 CC FT TRANSIT 1 32
 CC FT CHAIN 33 640 MITOCHONDRION (BY SIMILARITY).
 CC FT NP_BIND 255 262 GTP (POTENTIAL).
 CC FT ACT_SITE 306 306 BY SIMILARITY.
 CC FT CONFLICT 128 128 R -> C (IN REF. 1).
 CC FT CONFLICT 296 298 ALC -> RYV (IN REF. 3).
 CC FT CONFLICT 476 476 R -> S (IN REF. 3).
 CC FT SEQUENCE 640 AA; 70637 MW; B0DD2AB0B97B03D CRC64;
 SQ
 Query Match 92.3%; Score 36; DB 1; Length 640;
 Best Local Similarity 75.0%; Pred. No. 15;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EOEIRDLR 8
 DB 604 EOEIRDLR 611
 RESULT 8
 GFAP_MOUSE
 ID GFAP_MOUSE STANDARD; PRT; 430 AA.
 AC P03995;
 DT 23-OCT-1986 (Rel. 02, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glial fibrillary acidic protein, astrocyte (GFAP).
 GN GFAP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-8529756; PubMed-2994002;
 RA Balcarak J.M., Cowan N.J.;
 RT "Structure of the mouse glial fibrillary acidic protein gene:
 RT implications for the evolution of the intermediate filament multigene
 RT family.";
 RT Nucleic Acids Res. 13:5527-5543(1985).

RN SEQUENCE OF 28-430 FROM N.A.
 RP MEDLINE-94194082; PubMed-6585825;
 RA Lewis S.A., Balcarek J.M., Krek V., Shelanski M., Cowan N.J.;
 RT "Sequence of a cDNA clone encoding mouse glial fibrillary acidic
 RT protein: structural conservation of intermediate filaments."
 RL Proc. Natl. Acad. Sci. U.S.A. 81:2743-2746(1984).
 RN [3]
 RP SEQUENCE OF 28-430 FROM N.A.
 RA MEDLINE-86101618; PubMed-3866511;
 RA Cowan N.J., Lewis S.A., Balcarek J.M., Krek V., Shelanski M.L.;
 RT "Structural implications of a cDNA clone encoding mouse glial
 RT fibrillary acidic protein."
 RL Ann. N.Y. Acad. Sci. 455:575-582(1985).
 RN [4]
 RP REVISIONS TO N-TERMINUS
 RA MEDLINE-90294716; PubMed-2163003;
 RA Brenner M., Lampel K., Nakatani Y., Mill J., Banner C., Mearow K.,
 RA Dohdawa M., Lipsky R., Freese E.;
 RT "Characterization of human cDNA and genomic clones for glial
 RT fibrillary acidic protein."
 RL Brain Res. Mol. Brain Res. 7:277-286(1990).
 RN [5]
 RP SEQUENCE OF 1-26 FROM N.A.
 RC TISSUE-Embryo;
 RA MEDLINE-95074269; PubMed-7983160;
 RA Ralston J.E., Lu X., Hutcheson A.M., Oulian R.A.;
 RT "Identification of two N-terminal non-alpha-helical domain motifs
 RT important in the assembly of glial fibrillary acidic protein."
 RL J. Cell Sci. 107:1935-1948(1994).
 CC -1- FUNCTION: GFAP, A CLASS-III INTERMEDIATE FILAMENT, IS A CELL-
 CC SPECIFIC MARKER THAT, DURING THE DEVELOPMENT OF THE CENTRAL
 CC NERVOUS SYSTEM, DISTINGUISHES ASTROCYTES FROM OTHER GLIAL CELLS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL: K01347; AAA37678.1; -
 DR EMBL: M25937; AAA37679.1; -
 DR EMBL: X02801; CAA26571.1; ALT_INIT.
 DR EMBL: X78141; CAA55020.1; -
 DR PIR: B60052; VEMSGF.
 DR MGD: MGI:95697; GFAP.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil.
 FT DOMAIN 1 69 HEAD.
 FT DOMAIN 374 374 ROD.
 FT DOMAIN 375 430 TAIL.
 FT DOMAIN 70 101 COIL 1A.
 FT DOMAIN 102 112 LINKER 1.
 FT DOMAIN 113 211 COIL 1B.
 FT DOMAIN 212 227 LINKER 12.
 FT DOMAIN 228 249 COIL 2A.
 FT DOMAIN 250 253 LINKER 2.
 FT DOMAIN 374 374 COIL 2B.
 FT DOMAIN 375 374 COIL 2B.
 FT CONFLICT 347 347 H -> Q (IN REF. 2 AND 3).
 SO SEQUENCE 430 AA; 49917 MW; 446554A65B82DC9 CRC64;
 Query Match 89.7%; Score 35; DB 1; Length 430;
 Best Local Similarity 75.0%; Pred. No. 16;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

RESULT 9
 ID XNIF_XENLA STANDARD; PRT; 470 AA.
 AC P35617;
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Low molecular weight neuronal intermediate filament (XNIF).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 453-463.
 RC TISSUE-Brain;
 RA MEDLINE-92356194; PubMed-1494944;
 RA Charnas L.R., Szabo B.G., Gainer H.;
 RT "Identification and developmental expression of a novel low molecular
 RT weight neuronal intermediate filament protein expressed in Xenopus
 RT laevis".
 RL J. Neurosci. 12:3010-3024(1992).
 CC -1- TISSUE SPECIFICITY: NERVOUS SYSTEM; IN AXONS IN THE PNS AND IN
 CC SMALL PERIKARYA IN THE DORSAL ROOT GANGLION.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN AXONS OF EARLY DIFFERENTIATING
 CC NEURONS AS WELL AS IN THE ADULT NERVOUS SYSTEM.
 CC -1- DOMAIN: THIS PROTEIN CONTAINS A CENTRAL ROD REGION CHARACTERISTIC
 CC OF IF PROTEINS BUT DISPLAYS DIVERGENT HEAD AND TAIL DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL: M86653; AAA83019.1; -
 DR PIR: A44841; A44841.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil.
 FT DOMAIN 1 91 HEAD.
 FT DOMAIN 92 395 ROD.
 FT DOMAIN 396 470 TAIL.
 FT DOMAIN 91 123 COIL 1A.
 FT DOMAIN 121 136 LINKER 1.
 FT DOMAIN 137 232 COIL 1B.
 FT DOMAIN 233 251 LINKER 12.
 FT DOMAIN 252 270 COIL 2A.
 FT DOMAIN 271 279 LINKER 2.
 FT DOMAIN 280 395 COIL 2B.
 SO SEQUENCE 470 AA; 53672 MW; CB462E2B6BAC41C CRC64;
 Query Match 89.7%; Score 35; DB 1; Length 470;
 Best Local Similarity 87.5%; Pred. No. 17;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 10
 ID NFL_COTJA STANDARD; PRT; 555 AA.
 AC 002916;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-Oct-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (Neurofilament light polypeptide) (NF-L).
 DE L).
 GN NEFL.
 OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Coturnix.
 CC Coturnix.
 CC NCBI_TaxID=993934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Brain;
 RA MEDLINE=93224534; PubMed=8468353;
 RA Ohara O., Gahara Y., Miyake T., Teraoka H., Kitamura T.;
 RT "Neurofilament deficiency in quail caused by nonsense mutation in neurofilament-L gene.";
 RT Cell Biol. 121:387-395(1993).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M, AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- DISEASE: NF-L DEFICIENCY CAUSES THE DISORDER OUIVER.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to license@sdb-sdb.ch).
 CC -----
 CC EMBL: D13223; BAA02504.1; -
 CC EMBL: D13222; BAA02503.1; ALT_TERM.
 CC InterPro: IPR001664; IF.
 CC Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF, 1.
 DR KMW Intermediate filament; Coiled coil; Neurone.
 FT INIT_MET 0
 FT DOMAIN 1 93
 FT DOMAIN 94 397
 FT DOMAIN 398 555
 FT DOMAIN 94 125
 FT DOMAIN 126 138
 FT DOMAIN 139 234
 FT DOMAIN 235 253
 FT DOMAIN 254 272
 FT DOMAIN 273 281
 FT DOMAIN 282 397
 FT DOMAIN 398 444
 FT DOMAIN 445 555
 FT SEQUENCE 555 AA; 62282 MW; 98957ABD8EBA7712 CRC64;
 SQ
 Query Match 89.7%; Score 35; DB 1; Length 555;
 Best Local Similarity 87.5%; Pred. No. 20;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DT 01-APR-1993 (Rel. 25, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Pyruvate kinase, isozymes R/L (EC 2.7.1.40) (R-type/L-type pyruvate kinase) (Red cell/Liver pyruvate kinase).
 DE PKLR OR PKL OR PKL.
 GN PKLR OR PKL OR PKL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT TOKYO.
 RC MEDLINE=91376115; PubMed=1896471;
 RA Kanno H., Fujii H., Hirose A., Miwa S.;
 RT "cDNA cloning of human R-type pyruvate kinase and identification of a single amino acid substitution (Thr384->Met) affecting enzymatic stability in a pyruvate kinase variant (PK Tokyo) associated with hereditary hemolytic anemia.";
 RT Proc. Natl. Acad. Sci. U.S.A. 88:8218-8221(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=88158079; PubMed=3126495;
 RA Tani K., Fujii H., Nagata S., Miwa S.;
 RT "Human liver type pyruvate kinase: complete amino acid sequence and the expression in mammalian cells.";
 RT Proc. Natl. Acad. Sci. U.S.A. 85:1792-1795(1988).
 RN [3]
 RP REVISIONS TO 130 AND 232.
 RA Kanno H.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 470-574 FROM N.A.
 RC MEDLINE=87184521; PubMed=3566732;
 RA Tani K., Fujii H., Tsutsumi H., Sukeyama J., Toyoshima K., Yoshida M.C., Noguchi T., Tanaka T., Miwa S.;
 RT "Human liver type pyruvate kinase: cDNA cloning and chromosomal assignment.";
 RT Biochem. Biophys. Res. Commun. 143:431-438(1987).
 RN [5]
 RP SEQUENCE OF 365-431 FROM N.A., AND VARIANT OSAKA.
 RC MEDLINE=93236593; PubMed=8476433;
 RA Kanno H., Fujii H., Tsujino G., Miwa S.;
 RT "Molecular basis of impaired pyruvate kinase isozyme conversion in erythroid cells: a single amino acid substitution near the active site and decreased mRNA content of the R-type PK.";
 RT Biochem. Biophys. Res. Commun. 193:46-52(1993).
 RN [6]
 RP REVIEW ON VARIANTS.
 RC MEDLINE=96263732; PubMed=8664896;
 RA Beutler E., Baronciani L.;
 RT "Mutations in pyruvate kinase.";
 RT Hum. Mutat. 7:1-6(1996).
 RN [7]
 RP REVIEW ON VARIANTS.
 RC MEDLINE=96400713; PubMed=8807089;
 RA Baronciani L., Bianchi P., Zanella A.;
 RT "Hematologically important mutations: red cell pyruvate kinase.";
 RT Blood Cells Mol. Dis. 22:85-89(1996).
 RN [8]
 RP REVIEW ON VARIANTS.
 RC MEDLINE=97230013; PubMed=9075576;
 RA Baronciani L., Bianchi P., Zanella A.;
 RT "Hematologically important mutations: red cell pyruvate kinase (1st update).";
 RT Blood Cells Mol. Dis. 22:259-264(1996).
 RN [9]
 RP REVIEW ON VARIANTS.
 RC MEDLINE=99187977; PubMed=10087985;
 RA Baronciani L., Bianchi P., Zanella A.;
 RT "Hematologically important mutations: red cell pyruvate kinase (2nd update).";
 RT Blood Cells Mol. Dis. 24:273-279(1998).
 RN [10]

RP REVIEW ON VARIANTS.
 RX MEDLINE-20238076; PubMed-10772876;
 RA Bianchi P., Zanello A.;
 RT "Hemotologically important mutations: red cell pyruvate kinase (third
 update).";
 RL Blood Cells Mol. Dis. 26:47-53(2000).
 [11]
 RP VARIANTS CNSHA LINZ CYS-163 AND TOKYO/BEIRUT MET-384.
 RX MEDLINE-91208396; PubMed-2018831;
 RA Neubauer B., Lakomek M., Minkler H., Parke M., Hofferbert S.,
 Schrotter W.;
 RT "Point mutations in the L-type pyruvate kinase gene of two children
 with hemolytic anemia caused by pyruvate kinase deficiency.";
 RL Blood 77:1871-1875(1991).
 [12]
 RP VARIANTS CNSHA FUKUSHIMA/NAEBAISHI LYS-421.
 RX MEDLINE-92163106; PubMed-1536957;
 RA Kanno H., Fujii H., Hiroo A., Omine M., Miwa S.;
 RT "Identical point mutations of the R-type pyruvate kinase (PK) CDNA
 found in unrelated PK variants associated with hereditary hemolytic
 anemia.";
 RL Blood 79:1347-1350(1992).
 [13]
 RP VARIANTS CNSHA SAPPORO GLN-426.
 RX MEDLINE-93244440; PubMed-8481523;
 RA Kanno H., Fujii H., Miwa S.;
 RT "Low substrate affinity of pyruvate kinase variant (PK Sapporo)
 caused by a single amino acid substitution (426 Arg->Gln) associated
 with hereditary hemolytic anemia.";
 RL Blood 81:2439-2441(1993).
 [14]
 RP VARIANTS CNSHA D-134; P-155; H-359; W-486; V-495 AND Q-510.
 RX MEDLINE-93248282; PubMed-8483951;
 RA Baronciani L., Beutler E.;
 RT "Analysis of pyruvate kinase-deficiency mutations that produce
 nonspherocytic hemolytic anemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:4324-4327(1993).
 [15]
 RP VARIANTS CNSHA AMISH HIS-479.
 RX MEDLINE-94214145; PubMed-8161798;
 RA Kanno H., Balas S.K., Miwa S., Fujii H., Bowman H.S.;
 RT "Molecular abnormality of erythrocyte pyruvate kinase deficiency in
 the Amish.";
 RL Blood 83:2311-2316(1994).
 [16]
 RP VARIANTS S-332; S-336; K-334 DEL; D-361; T-392; H-498; Q-510 AND
 W-532.
 RX MEDLINE-94235845; PubMed-8180378;
 RA Lennner C., Nuerberg P., Thiele B.-J., Reis A., Brabec V.,
 Sakalova A., Jacobasch G.;
 RT "Mutations in the pyruvate kinase L gene in patients with hereditary
 hemolytic anemia.";
 RL Blood 83:2817-2822(1994).
 [17]
 RP VARIANTS CNSHA E-331; A-341; K-393; S-393; D-458; M-460 AND H-498.
 RX MEDLINE-95221622; PubMed-7706479;
 RA Baronciani L., Beutler E.;
 RT "Molecular study of pyruvate kinase deficient patients with
 hereditary nonspherocytic hemolytic anemia.";
 RL J. Clin. Invest. 95:1702-1709(1995).
 [18]
 RP VARIANTS CNSHA.
 RA Baronciani L., Westwood B., Beutler E.;
 RT "Study of the molecular defects in pyruvate kinase (PK) deficient
 patients affected by hereditary nonspherocytic hemolytic anemia
 (HNHA).";
 RL J. Invest. Med. 43:341A-341A(1995).
 [19]
 RP VARIANTS CNSHA GLU-37.
 RX MEDLINE-97245895; PubMed-9090535;
 RA Beutler E., Westwood B., van Zylsten R., Roos D.;
 RT "G-to-T transition at CDNA nt 110 (K37Q) in the PKLR (pyruvate
 kinase) gene is the molecular basis of a case of hereditary increase

RT of red blood cell ATP.";
 RL Hum. Mutat. 9:282-285(1997).
 [20]
 RP VARIANTS O-172; Q-337; H-339; T-357; I-408; T-431; W-486 AND Q-532.
 RX MEDLINE-99043610; PubMed-9827908;
 RA Zarza R., Alvarez R., Pujades A., Nondedeu B., Carreras A., Estella J.,
 Remacha A., Sanchez J.M., Morey M., Cortes T., Perez Lungmus G.,
 Bureo E., Vives Corrons J.L.;
 RT "Molecular characterization of the PK-LR gene in pyruvate kinase
 deficient Spanish patients.";
 RL Br. J. Haematol. 103:377-382(1998).
 [21]
 RP VARIANTS CNSHA CONAKRY TYR-130.
 RX MEDLINE-99101396; PubMed-9886305;
 RA Cohen-Solal M., Prehu C., Majcman H., Poyart C.,
 Galacteros F.;
 RT "A new sickle cell disease phenotype associating Hb S trait, severe
 pyruvate kinase deficiency (PK Conakry), and an alpha-2 globin gene
 variant (Hb Conakry).";
 RL Br. J. Haematol. 103:950-956(1998).
 [22]
 RP VARIANTS CNSHA S-332; P-337; W-486; C-498 AND Q-510.
 RX MEDLINE-9811680; PubMed-9482576;
 RA Pastore L., della Morte R., Frisio G., Affilito F., Vitale D.,
 Calise R.M., Ferraro F., Zagari A., Rocco B., Salvatore F.;
 RT "Novel mutations and structural implications in R-type pyruvate
 kinase-deficient patients from Southern Italy.";
 RL Hum. Mutat. 11:127-134(1998).
 [23]
 RP VARIANTS CNSHA M-335; K-338 DEL; G-387; D-394 AND V-394.
 RX MEDLINE-2126334; PubMed-1138279;
 RA Zanello A., Bianchi P., Fermo E., Iurlo A., Zappa M., Vercellatti C.,
 Boschetti C., Baronciani L., Cotton F.;
 RT "Molecular characterization of the PK-LR gene in sixteen pyruvate
 kinase-deficient patients.";
 RL Br. J. Haematol. 113:43-48(2001).
 CC -1- CATALYTIC ACTIVITY: ATP + Pyruvate -> ADP + phosphoenolpyruvate.
 CC -1- COFACTOR: REQUIRES MAGNESIUM AND POTASSIUM.
 CC -1- PATHWAY: GLYCOLYSIS; final step.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; L-TYPE AND R-TYPE (SHOWN HERE);
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- DISEASE: DEFECTS IN PKLR ARE THE MOST COMMON CAUSE OF CHRONIC
 CC HEREDITARY NONSPHEROCYTIC HEMOLYTIC ANEMIA (CNSHA OR HNHA) AMONG
 CC GLYCOLYTIC ENZYMES.
 CC -1- MISCELLANEOUS: THERE ARE 4 ISOZYMES OF PYRUVATE KINASE IN MAMMALS:
 CC L, R, M1 AND M2. L TYPE IS MAJOR ISOZYME IN THE LIVER, R IS FOUND
 CC IN RED CELLS, M1 IS THE MAIN FORM IN MUSCLE, HEART AND BRAIN, AND
 CC M2 IS FOUND IN EARLY FETAL TISSUES.
 CC -1- SIMILARITY: BELONGS TO THE PYRUVATE KINASE FAMILY.
 CC -----
 Query Match 89.7%; Score 35; DB 1; Length 574;
 Best Local Similarity 75.0%; Pred. No. 21;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EOEIRDLR 8
 DB 266 EODVRLDR 273
 RESULT 12
 PIR4_BOVIN STANDARD; PRT; 1023 AA.
 ID Q07722; Q07721; Q07723; Q07724;
 AC 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase beta 4
 DE (EC 3.1.4.11) (PLC-beta-4) (phospholipase C-beta-4) (Fragment).
 GN PLCB4.
 OS Bos taurus (Bovine).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Kuminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_TaxID=9913;
 CC [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE-RETINA;
 RX MEDLINE=93317620; PubMed-8327481;
 RA Ferreira P.A., Shortridge R.D., Pak W.L.;
 RT "Distinctive subtypes of bovine phospholipase C that have preferential
 expression in the retina and high homology to the norpa gene product
 of *Drosophila*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6042-6046(1993).
 CC -1- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES
 CC DIACYLGLYCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS
 CC MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE
 CC C ENZYMES. THIS FORM HAS A ROLE IN RETINA SIGNAL TRANSDUCTION.
 CC -1- CATALYTIC ACTIVITY: 1-phosphatidyl-1D-myo-inositol 4,5-
 CC bisphosphate + H(2)O -> D-myo-inositol 1,4,5-trisphosphate +
 CC diacylglycerol.
 CC -1- COFACTOR: Calcium.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS: 1A, 1B (SHOWN HERE), 2A
 CC AND 2B, ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE RETINA.
 CC -1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN FOUR DIFFERENT FORMS
 CC OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L13936; AAA30699.1; -
 DR EMBL: L13935; AAC37304.1; -
 DR EMBL: L13937; AAA30700.1; -
 DR EMBL: L13938; AAA30701.1; -
 DR HSSP: P10688; 10AS.
 DR InterPro: IPR000008; C2.
 DR InterPro: IPR001192; PI_PLC.
 DR InterPro: IPR000909; PI_PLC_Xdom.
 DR InterPro: IPR001711; PI_PLC_X.
 DR Pfam: PF00387; PI-PLC-Y; 1.
 DR Pfam: PF00388; PI-PLC-X; 1.
 DR PRINTS: PR00390; PPHPLIPASEC.
 DR PRODOM: PD001202; PI_PLC_Y; 1.
 DR SMART: SM00339; C2; 1.
 DR SMART: SM00148; PLCXc; 1.
 DR SMART: SM00149; PLCYc; 1.
 DR PROSITE: PS50004; C2_DOMAIN_2; 1.
 DR PROSITE: PS50007; PIPLC_XDOMAIN; 1.
 DR PROSITE: PS50008; PIPLC_YDOMAIN; 1.
 KW Hydrolyase; Lipid degradation; Transducer; Phosphorylation; Calcium;
 KW Alternative splicing.
 FT DOMAIN 149 299 DOMAIN X.
 FT DOMAIN 413 529 DOMAIN Y.
 FT DOMAIN 536 634 C2 DOMAIN.
 FT ACT_SITE 164 164 BY SIMILARITY.
 FT ACT_SITE 211 211 BY SIMILARITY.
 FT VARSPLIC 372 383 MISSING (IN ISOFORM 1A).
 FT VARSPLIC 372 383 MISSING (IN ISOFORM 2A).
 SO SEQUENCE 1023 AA; 117109 MW; 416D700C2095748C CRC64;

Query Match 87.2%; Score 34; DB 1; Length 1023;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 EOEIRDL 7
 Db 883 EOEIRDL 889
 RESULT 13
 ID PIB4.HUMAN STANDARD; PRT; 1175 AA.
 AC Q15147; O9BQW5; O9BQW6; O9BQW8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase beta 4
 DE (EC 3.1.4.11) (PLC-beta-4) (Phospholipase C-beta-4).
 GN PLCB4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxID=9606;
 CC [1]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE-RETINA;
 RX MEDLINE=96079091; PubMed-8530101;
 RA Alvarez R.A., Ghalayini A.J., Xu P., Hardcastle A., Bhattacharya S.,
 RA Rao P.N., Pettenati M.V., Anderson R.E., Baehr W.;
 RT "cDNA sequence and gene locus of the human retinal phospholipase-
 RT specific phospholipase-C beta 4 (PLCB4).";
 RL Genomics 29:53-61(1995).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3).
 RP MEDLINE=21638749; PubMed-11780052;
 RX DeLoulas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggaley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cory N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahman D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knigths A., Laird G.K., Lawlor S.,
 RA Lehaszaiho M.H., Leverisha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachle L.J., McLeay K., Murray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prithalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -1- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES
 CC DIACYLGLYCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS
 CC MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE
 CC C ENZYMES. THIS FORM HAS A ROLE IN RETINA SIGNAL TRANSDUCTION.
 CC -1- CATALYTIC ACTIVITY: 1-phosphatidyl-1D-myo-inositol 4,5-
 CC bisphosphate + H(2)O -> D-myo-inositol 1,4,5-trisphosphate +
 CC diacylglycerol.
 CC -1- COFACTOR: Calcium.
 CC -1- ALTERNATIVE PRODUCTS: At least 3 isoforms: 1, 2 (shown here) and
 CC 3, are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE RETINA.
 CC -1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN FOUR DIFFERENT FORMS
 CC OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AL031652; CAC21068.1; -
 DR EMBL: AL023805; CAC34525.1; -
 DR EMBL: AL023805; CAC34527.1; -
 DR EMBL: AL023805; CAC34528.1; -
 DR EMBL: LA1349; AAB02027.1; -
 DR HSSP: P10688; IDJX.
 DR Genew: HGNC:9059; PLCB4.
 DR MIM: 600810; -
 DR Interpro: IPR000008; C2.
 DR Interpro: IPR001192; PI_PLC.
 DR Interpro: IPR000909; PI_PLC_xdom.
 DR Interpro: IPR001711; PI_PLC_Y.
 DR Pfam: PF00168; C2; 1.
 DR Pfam: PF00387; PI_PLC-X; 1.
 DR Pfam: PF00388; PI_PLC-X; 3.
 DR PRINTS: PR00390; PPHPLIPASEC.
 DR PRODOM: PD001202; PI_PLC_Y; 1.
 DR SMART: SM00239; C2; 1.
 DR SMART: SM00148; PLCXC; 1.
 DR SMART: SM00149; PLCXC; 1.
 DR PROSITE: PSS0004; C2_DOMAIN_2; 1.
 DR PROSITE: PSS0007; PIPLC_X_DOMAIN; 1.
 DR PROSITE: PSS0008; PIPLC_Y_DOMAIN; 1.
 DR Hydrolase: Lipid degradation; Transducer; Phosphorylation; Calcium;
 KM Alternative splicing.
 FT DOMAIN 313 463 DOMAIN X.
 FT DOMAIN 565 681 DOMAIN Y.
 FT ACT_SITE 688 786 C2 DOMAIN.
 FT ACT_SITE 328 328 BY SIMILARITY.
 FT ACT_SITE 375 375 BY SIMILARITY.
 FT VARSPIC 1 153 MISSING (IN ISOFORM 1).
 FT VARSPIC 154 167 LAFWNTNGKIPVR -> MNMNMVCFELCP
 (IN ISOFORM 1).
 FT VARSPIC 1 268 MISSING (IN ISOFORM 3).
 FT CONFLICT 447 447 A -> P (IN REF. 1).
 FT CONFLICT 757 757 F -> L (IN REF. 1).
 FT CONFLICT 787 787 L -> P (IN REF. 1).
 FT CONFLICT 840 840 K -> T (IN REF. 1).
 FT CONFLICT 902 902 A -> P (IN REF. 1).
 SQ SEQUENCE 1175 AA; 134463 MW; AB2C8EB99EF57357 CRC64;

Query Match 87.28; Score 34; DB 1; Length 1175;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDL 7
 DB 1035 EOEIRDL 1041

RESULT 14
 PIB4_RAT STANDARD; PRT: 1175 AA.
 AC Q9QW07; Q92066; O88356;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase beta 4
 DE (EC 3.1.4.11) (PLC-beta-4) (Phospholipase C-beta-4).
 GN PLCB4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_Taxid=10116;
 RN [1]

RP SEQUENCE FROM N.A. (ISOFORM A).
 RC STRAIN-Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=94012687; PubMed=8407970;
 RA Lee C.-W., Park D.J., Lee K.-H., Kim C.G., Rhee S.G.;
 RT "Purification, molecular cloning, and sequencing of phospholipase C-
 RT beta 4." J. Biol. Chem. 268:21318-21327(1993).
 RL J. Biol. Chem. 268:21318-21327(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM B).
 RC STRAIN-Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=93343926; PubMed=7688223;
 RA Kim M.J., Baik Y.Y., Min D.S., Lee S.J., Ryu S.H., Suh P.G.;
 RT "Cloning of cDNA encoding rat phospholipase C-beta 4, a new member of
 RT the phospholipase C." Biochem. Biophys. Res. Commun. 194:706-712(1993).
 RL Biochem. Biophys. Res. Commun. 194:706-712(1993).
 RN [3]
 RP SEQUENCE OF 447-1175 FROM N.A. (ISOFORM C).
 RC TISSUE=Brain;
 RX MEDLINE=99132015; PubMed=9931434;
 RA Adamaki F.M., Timms K.M., Shieh B.H.;
 RT "A unique isoform of phospholipase Cbeta4 highly expressed in the
 RT cerebellum and eye." Biochim. Biophys. Acta 1444:55-60(1999).
 RL Biochim. Biophys. Acta 1444:55-60(1999).
 CC -1- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES
 CC DIACYLGLYCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS
 CC MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE
 CC C ENZYMES. THIS FORM HAS A ROLE IN RETINA SIGNAL TRANSDUCTION.
 CC -1- CATALYTIC ACTIVITY: 1-phosphatidyl-1D-myo-inositol 4,5-
 CC bisphosphate + H(2)O -> D-myo-inositol 1,4,5-trisphosphate +
 CC diacylglycerol.
 CC -1- COFACTOR: Calcium.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A (SHOWN HERE), B AND C; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN THE RETINA.
 CC -1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN FOUR DIFFERENT FORMS
 CC OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L15556; AAK1357.1; -
 DR EMBL: U57836; AAD10403.1; -
 DR EMBL: AF031370; AAC98145.1; -
 DR EMBL: AF027571; AAC24984.1; -
 DR HSSP: P10688; IDJX.
 DR Interpro: IPR000008; C2.
 DR Interpro: IPR001192; PI_PLC.
 DR Interpro: IPR000909; PI_PLC_xdom.
 DR Interpro: IPR001711; PI_PLC_Y.
 DR Pfam: PF00168; C2; 1.
 DR Pfam: PF00387; PI_PLC-X; 1.
 DR Pfam: PF00388; PI_PLC-X; 3.
 DR PRINTS: PR00390; PPHPLIPASEC.
 DR PRODOM: PD001202; PI_PLC_Y; 1.
 DR SMART: SM00239; C2; 1.
 DR SMART: SM00148; PLCXC; 1.
 DR SMART: SM00149; PLCXC; 1.
 DR PROSITE: PSS0004; C2_DOMAIN_2; 1.
 DR PROSITE: PSS0007; PIPLC_X_DOMAIN; 1.
 DR PROSITE: PSS0008; PIPLC_Y_DOMAIN; 1.
 DR Hydrolase: Lipid degradation; Transducer; Phosphorylation; Calcium;
 KM Alternative splicing.
 FT DOMAIN 313 463 DOMAIN X.
 FT DOMAIN 565 681 DOMAIN Y.
 FT ACT_SITE 688 786 C2 DOMAIN.
 FT ACT_SITE 328 328 BY SIMILARITY.
 FT ACT_SITE 375 375 BY SIMILARITY.

```

FT VARSPLIC 1013 1022 VKETVAGHTK -> GKORDASPSG (IN ISOFORM B).
FT VARSPLIC 1023 1175 MISSING (IN ISOFORM B).
FT VARSPLIC 1154 1175 AKEMQWVLEAEMDRATV -> LKSCHAVSOTQGES
DAAGETISRDGPOTSNSMILQAN (IN ISOFORM
C).
FT CONFLICT 255 255 L -> M (IN REF. 2).
FT CONFLICT 308 308 R -> A (IN REF. 2).
FT CONFLICT 417 417 Q -> E (IN REF. 2).
FT CONFLICT 470 470 E -> K (IN REF. 2).
FT CONFLICT 504 504 A -> AA (IN REF. 1).
FT CONFLICT 545 546 EQ -> DE (IN REF. 2).
FT CONFLICT 734 734 I -> L (IN REF. 2).
FT CONFLICT 741 741 R -> H (IN REF. 2).
FT CONFLICT 764 764 L -> M (IN REF. 2).
FT CONFLICT 776 776 D -> N (IN REF. 2).
FT CONFLICT 828 828 F -> L (IN REF. 1).
FT CONFLICT 843 843 S -> Y (IN REF. 2).
FT CONFLICT 852 852 L -> M (IN REF. 3).
FT CONFLICT 916 916 Q -> T (IN REF. 2).
FT CONFLICT 1024 1024 W -> C (IN REF. 3).
FT CONFLICT 1043 1043 L -> M (IN REF. 3).
FT CONFLICT 1057 1057 A -> V (IN REF. 3).
FT CONFLICT 1067 1067 L -> V (IN REF. 3).
FT CONFLICT 1084 1084 S -> C (IN REF. 3).
SQ SEQUENCE 1175 AA; 134496 MM; 73796BB95B8FCD CRC64;

```

```

Query Match
Best Local Similarity 100.0%; Score 34; DB 1; Length 1175;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 EQEIRDL 7
DB 1035 EQEIRDL 1041

```

```

RESULT 15
IF3L_TORCA STANDARD; PRT; 458 AA.
AC P23729;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Type III intermediate filament.
OS Torpedo californica (Pacific electric ray).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalea; Hypnosqualea; Pristiogalea; Batoidae;
OC Torpediniformes; Torpedinidae; Torpedo.
OX NCBI_TaxID=7787;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90245604; PubMed=2336376;
RA Frail D.E., Mudd J., Merlie J.P.;
RT "Nucleotide sequence of an intermediate filament cDNA from Torpedo
RT californica."
RL Nucleic Acids Res. 18:1910-1910(1990).
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X51533; CA35912.1; -
CC PIR: S09228; S09228.
CC InterPro: IPR001664; IF.
CC Pfam: PF00038; filament; 1.
CC PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil.
FT DOMAIN 1 100 HEAD.
FT DOMAIN 101 405 ROD.

```

```

FT DOMAIN 406 458 TAIL.
SQ SEQUENCE 458 AA; 52819 MM; F30B0B235A6644FD CRC64;
Query Match
Best Local Similarity 84.6%; Score 33; DB 1; Length 458;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 EQEIRDL 8
DB 145 EQEIRDL 152

```

```

Search completed: November 13, 2002, 13:16:55
Job time : 7.6383 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 25.7021 Seconds
(Without alignments)
64.134 Million cell updates/sec

Title: US-09-856-086-7

Perfect score: 39

Sequence: 1 EOEIRDLR 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.21.*

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_ricent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioplasmid:*
17: sp_archaeoplasmid:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	100.0	386	4	Q8TCR7
2	36	92.3	198	12	Q9WQX8
3	36	92.3	199	12	Q9YRV6
4	36	92.3	199	12	Q9WQY4
5	36	92.3	199	12	Q86365
6	36	92.3	199	12	Q8V6W2
7	36	92.3	222	12	Q80303
8	36	92.3	453	13	Q73587
9	36	92.3	535	11	Q8R3X7
10	36	92.3	640	11	Q91210
11	36	92.3	1380	10	Q9FJ57
12	35	89.7	422	5	Q96755
13	35	89.7	566	4	Q16716
14	35	89.7	587	4	Q16715
15	35	89.7	599	4	Q75758
16	35	89.7	1059	5	Q960T6

17	34	87.2	231	16	Q9KBM6	Q9KBM6 bacillus ha
18	34	87.2	467	3	Q96VC3	Q96VC3 mucor circi
19	34	87.2	497	3	Q9P3C2	Q9P3C2 neurospora
20	34	87.2	605	16	Q8Y270	Q8Y270 talstonia s
21	34	87.2	781	16	Q8Y0R5	Q8Y0R5 talstonia s
22	34	87.2	1175	11	Q91U21	Q91U21 mus musculu
23	34	87.2	1975	5	Q9VCD1	Q9VCD1 drosophila
24	33	84.6	84	15	Q91182	Q91182 human limun
25	33	84.6	85	15	Q99DA2	Q99DA2 human limun
26	33	84.6	85	15	Q99D92	Q99D92 human limun
27	33	84.6	85	15	Q9QND9	Q9QND9 human limun
28	33	84.6	85	15	Q91142	Q91142 human limun
29	33	84.6	85	15	Q91154	Q91154 human limun
30	33	84.6	85	15	Q91160	Q91160 human limun
31	33	84.6	85	15	Q91178	Q91178 human limun
32	33	84.6	85	15	Q903R4	Q903R4 human limun
33	33	84.6	85	15	Q900Z5	Q900Z5 human limun
34	33	84.6	85	15	Q900Y6	Q900Y6 human limun
35	33	84.6	199	12	Q9YRV9	Q9YRV9 rice tungro
36	33	84.6	386	5	Q62003	Q62003 branchiosto
37	33	84.6	415	9	Q80046	Q80046 staphylococ
38	33	84.6	415	9	Q8SDK8	Q8SDK8 staphylococ
39	33	84.6	435	5	Q95SS5	Q95SS5 drosophila
40	33	84.6	450	13	Q9DDB3	Q9DDB3 scyllorhinu
41	33	84.6	455	5	Q9V8D7	Q9V8D7 drosophila
42	33	84.6	914	10	Q24371	Q24371 solanum tub
43	33	84.6	1212	5	Q9XZ29	Q9XZ29 drosophila
44	33	84.6	1678	5	Q9NFE6	Q9NFE6 caenorhabdi
45	33	84.6	2288	5	Q23081	Q23081 caenorhabdi

ALIGNMENTS

RESULT 1

ID	Q8TCR7	PRELIMINARY;	PRT;	386 AA.
AC	Q8TCR7			
DT	01-JUN-2002 (TREMREL. 21, Created)			
DT	01-JUN-2002 (TREMREL. 21, Last sequence update)			
DT	01-JUN-2002 (TREMREL. 21, Last annotation update)			
DE	Hypothetical 43.8 kDa protein (Fragment).			
GN	DKEZP761K0922.			
OS	Human sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=AMGDALA;			
RA	Wambutt R., Heubner D., Meyers H.W., Well B., Wiemann S.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL, AL713644; CAD28456.1; ..			
KW	Hypothetical protein.			
FT	NON_TER			
SQ	SEQUENCE 386 AA; 43779 MW; F293388B200C7B65 CRC64;			

Query Match	100.0%;	Score 39;	DB 4;	Length 386;
Best Local Similarity	100.0%;	Pred. NO. 12;		
Matches	8;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

QY 1 EOEIRDLR 8
|||||||

DB 27 EOEIRDLR 34

RESULT 2

ID	Q9WQX8	PRELIMINARY;	PRT;	198 AA.
AC	Q9WQX8			
DT	01-NOV-1999 (TREMREL. 12, Created)			
DT	01-NOV-1999 (TREMREL. 12, Last sequence update)			
DT	01-NOV-1999 (TREMREL. 12, Last annotation update)			

```

DE P24 (Fragment).
OS Rice tungro bacilliform virus.
OC Viruses; Retroviral viruses; Caulimoviridae;
OC Rice tungro bacilliform-like viruses.
OX NCBI_TaxID=10654;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RTBV-1C;
RX MEDLINE=99394699; PubMed=10466823;
RA Cabauatan P.O., Melcher U., Ishikawa K., Omura T., Hibino H.,
RA Koganezawa H., Azam O.;
RT "Sequence changes in six variants of rice tungro bacilliform virus and
RT their phylogenetic relationships.";
RL J. Gen. Virol. 80:2229-2237(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RTBV-1C;
RA Cabauatan P.O., Melcher U., Ishikawa K., Omura T., Hibino H.,
RA Koganezawa H., Azam O.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF113832; AAD30196.1; -.
FT NON_TER 1
SQ SEQUENCE 198 AA; 23595 MW; 26CB578019CE957D CRC64;

Query Match 92.3%; Score 36; DB 12; Length 199;
Best Local Similarity 87.5%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
DB 180 EOEIKDLR 187

RESULT 3
OYRV6 PRELIMINARY; PRT; 199 AA.
ID OYRV6;
AC OYRV6;
DT 01-NOV-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE P24 (Fragment).
OS Rice tungro bacilliform virus.
OC Viruses; Retroviral viruses; Caulimoviridae;
OC Rice tungro bacilliform-like viruses.
OX NCBI_TaxID=10654;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RTBV-1C;
RA Druka A., Chetanachit D., Rattanakarn W., Natwong B., Hay J.,
RA Burns T., Dischoporn S., Hull R.;
RT "Comparative analysis of structural proteins of two rice tungro
RT bacilliform virus isolates from Thailand.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF094571; AAC9861.1; -.
FT NON_TER 1
SQ SEQUENCE 199 AA; 23682 MW; 827B7B63CE83E45 CRC64;

Query Match 92.3%; Score 36; DB 12; Length 199;
Best Local Similarity 87.5%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
DB 181 EOEIKDLR 188

RESULT 4
OYRV4 PRELIMINARY; PRT; 199 AA.
ID OYRV4;
AC OYRV4;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

```

```

DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE P24 (Fragment).
OS Rice tungro bacilliform virus.
OC Viruses; Retroviral viruses; Caulimoviridae;
OC Rice tungro bacilliform-like viruses.
OX NCBI_TaxID=10654;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RTBV-G1;
RX MEDLINE=99394699; PubMed=10466823;
RA Cabauatan P.O., Melcher U., Ishikawa K., Omura T., Hibino H.,
RA Koganezawa H., Azam O.;
RT "Sequence changes in six variants of rice tungro bacilliform virus and
RT their phylogenetic relationships.";
RL J. Gen. Virol. 80:2229-2237(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RTBV-G1;
RA Cabauatan P.O., Melcher U., Ishikawa K., Omura T., Hibino H.,
RA Koganezawa H., Azam O.; the EMBL/GenBank/DBJ databases.
DR EMBL; AF113830; AAD30188.1; -.
FT NON_TER 1
SQ SEQUENCE 199 AA; 23723 MW; 96A3482082A2B19A CRC64;

Query Match 92.3%; Score 36; DB 12; Length 199;
Best Local Similarity 87.5%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
DB 181 EOEIKDLR 188

RESULT 5
OYRV5 PRELIMINARY; PRT; 199 AA.
ID OYRV5;
AC OYRV5;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-1999 (TREMBlrel. 09, Last annotation update)
DE 24k polypeptide (Fragment).
OS Rice tungro bacilliform virus.
OC Viruses; Retroviral viruses; Caulimoviridae;
OC Rice tungro bacilliform-like viruses.
OX NCBI_TaxID=10654;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92238849; PubMed=1571015;
RA Keno H., Koizumi M., Noda H., Hibino H., Ishikawa K., Omura T.,
RA Cabauatan P.O., Koganezawa H.;
RT "Nucleotide sequence of capsid protein gene of rice tungro bacilliform
RT virus.";
RL Arch. Virol. 124:157-163(1992).
DR EMBL; D10774; BAA01605.1; -.
FT NON_TER 1
SQ SEQUENCE 199 AA; 23636 MW; 14B38E664E815D6 CRC64;

Query Match 92.3%; Score 36; DB 12; Length 199;
Best Local Similarity 87.5%; Pred. No. 24;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
DB 181 EOEIKDLR 188

RESULT 6
OYRV2 PRELIMINARY; PRT; 199 AA.
ID OYRV2;
AC OYRV2;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)

```


01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE P24 (Fragment).
 OS Rice tungro bacilliform virus.
 CC Viruses; Retroviral viruses; Caulimoviridae;
 CC Rice tungro bacilliform-like viruses.
 CC NCBI_TaxID=10654;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CHAINAT ISOLATE;
 RA Nakhong B., Ratanakarn W., Chettanachit D.;
 RT "Complete Genomic Sequence of Rice Tungro Bacilliform Virus - Chainat
 RT isolate from Thailand."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF220561; AAL55649.1;
 FT NON_TER
 SQ SEQUENCE 199 AA; 23697 MW; 7C2485CFA7610DDC CRC64;
 SO

Query Match 92.3%; Score 36; DB 12; Length 199;
 Best Local Similarity 87.5%; Pred. No. 24;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 DB 181 EOEIRDLR 188

RESULT 7
 090303 PRELIMINARY; PRT; 222 AA.
 AC 090303;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical 26.5 kDa protein.
 OS Rice tungro bacilliform virus.
 CC Viruses; Retroviral viruses; Caulimoviridae;
 CC Rice tungro bacilliform-like viruses.
 CC NCBI_TaxID=10654;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SERDANG;
 RA MEDLINE=99119505; PubMed=9918990;
 RA Marney P., Botmer B., Jacquot E., de Kochko A., Ong C.A., Yot P.,
 RA Siuzdak G., Beachy R.N., Fauquet C.M.;
 RT "Rice tungro bacilliform virus open reading frame 3 encodes a single
 RT 37-kDa coat protein."
 RL Virology 253:319-326(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SERDANG;
 RA de Kochko A., Marney P., Brizard J.P., Beachy R.N., Fauquet C.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF076470; AAC27709.2;
 KW Hypothetical protein
 SQ SEQUENCE 222 AA; 26460 MW; 95DB05292B3760CA CRC64;
 SO

Query Match 92.3%; Score 36; DB 12; Length 222;
 Best Local Similarity 87.5%; Pred. No. 27;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 DB 204 EOEIRDLR 211

RESULT 8
 073587 PRELIMINARY; PRT; 455 AA.
 AC 073587;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Desmin.

GN DES.
 OS Scyllorhinus stellaris.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 CC Elasmobranchii; Galeomorphi; Galeidae; Carcharhiniformes;
 CC Scyllorhinidae; Scyllorhinus.
 CC NCBI_TaxID=68454;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MUSCLE;
 RA Schultess J., Loebbecke A., Schaffeld M., Lieb B., Mark J.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Schaffeld M., Schultess J., Loebbecke A., Lieb B., Herrmann H.,
 RA Mark J.;
 RT "Primary structure, expression patterns and properties of vimentin and
 RT desmin in the shark Scyllorhinus stellaris."
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 DR EMBL; Y15064; CAA75347.1;
 DR EMBL; AJ304374; CAC83054.1;
 DR InterPro: IPR001664; IF; 1.
 DR Pfam: PF00038; Filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Colled coll.; Intermediate filament.
 SQ SEQUENCE 455 AA; 52230 MW; B330A2FCF895BCE9 CRC64;
 SO

Query Match 92.3%; Score 36; DB 13; Length 455;
 Best Local Similarity 87.5%; Pred. No. 55;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 DB 140 EOEIRDLR 147

RESULT 9
 08R3X7 PRELIMINARY; PRT; 535 AA.
 AC 08R3X7;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Similar to RIKEN CDNA 9130022B02 gene (Fragment).
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC023437; AAH23437.1;
 FT NON_TER
 SQ SEQUENCE 535 AA; 58782 MW; 63664DC81D27C184 CRC64;
 SO

Query Match 92.3%; Score 36; DB 11; Length 535;
 Best Local Similarity 75.0%; Pred. No. 64;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 DB 499 EOEIRDLR 506

RESULT 10
 091210 PRELIMINARY; PRT; 640 AA.
 AC 091210;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Similar to phosphoenolpyruvate carboxykinase 2 (mitochondrial).

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Strausberg R.; (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC010318; AAH0318.1; -
 DR InterPro: IPR000364; PEP_carboxykin.
 DR Pfam: PF00821; PEPCK; 1.
 DR ProDom: PD004738; PEP_carboxykin; 1.
 DR PROSITE: PS00505; PEPCK_GTP; UNKNOWN_1.
 DR KINASE: PYruvate.
 KW KINASE: PYruvate.
 SQ SEQUENCE 640 AA; 70494 MW; F1CA0E783D197E60 CRC64;

Query Match
 Best Local Similarity 92.3%; Score 36; DB 11; Length 640;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 Db 604 EOEVRDLR 611

RESULT 11
 ID 09FJ57 PRELIMINARY; PRT; 1380 AA.
 AC 09FJ57;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE GB|AAD39572.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Euroside II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-COLUMBIA;
 RX MEDLINE-99087489; PubMed-9872454;
 RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
 RT Sequence features of the regions of 1,013,767 bp covered by sixteen
 RT physically assigned P1 and YAC clones.";
 RL DNA Res. 5:297-308(1998).
 DR EMBL; AB015479; BAB08571.1; -
 DR InterPro: IPR003592; LRR_out.
 DR SMART: SM00370; LRR; 5
 SQ SEQUENCE 1380 AA; 156946 MW; C853381B966CE3 CRC64;

Query Match
 Best Local Similarity 92.3%; Score 36; DB 10; Length 1380;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 Db 1141 EOEIRDLR 1148

RESULT 12
 ID 096755 PRELIMINARY; PRT; 422 AA.
 AC 096755;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Intermediate filament protein E1.
 OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
 OC Branchiostoma.

OX NCBI_TaxID=7740;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE-99019308; PubMed-9804163;
 RX Karabinos A., Riemer D., Erber A., Weber R.;
 RT "Homologues of vertebrate type I, II and III intermediate filament
 RT (IF) proteins in an invertebrate: the IF multigene family of the
 RT cephalochordate Branchiostoma";
 RL FEBS Lett. 437:15-18(1998).
 DR EMBL; AJ010294; CAA09068.1; -
 DR InterPro: IPR002952; Egshell.
 DR InterPro: IPR001664; IF.
 DR InterPro: IPR002957; Keratin_I.
 DR Pfam: PF00038; Egshell.
 DR PRINTS: PR01228; EGGSHELL.
 DR PRINTS: PR01248; TYPEI KERATIN.
 SQ SEQUENCE 422 AA; 44892 MW; 85FE742F0751B24 CRC64;

Query Match
 Best Local Similarity 89.7%; Score 35; DB 5; Length 422;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 Db 243 EOEVRDLR 250

RESULT 13
 ID 016716 PRELIMINARY; PRT; 566 AA.
 AC 016716;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Pyruvate kinase (EC 2.7.1.40) (PK).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE-98182587; PubMed-9522120;
 RX Lenzner C., Nurnberg P., Jacobasch G., Thiele B.J.;
 RT "Complete genomic sequence of the human PK-L/R-gene includes four
 RT intragenic polymorphisms defining different haplotype backgrounds of
 RT normal and mutant PK-genes";
 RL DNA Seq. 8:45-53(1997).
 CC -I- CATALYTIC ACTIVITY: ATP + PYRUVATE -> ADP + PHOSPHOENOLPYRUVATE.
 CC -I- COFACTOR: REQUIRES MAGNESIUM AND POTASSIUM (BY SIMILARITY).
 CC -I- PATHWAY: FINAL STEP IN GLYCOLYSIS.
 CC -I- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -I- SIMILARITY: BELONGS TO THE PYRUVATE KINASE FAMILY.
 DR EMBL; U47654; AAA92536.1; -
 DR HSSP: U47654; LAOF.
 DR InterPro: IPR000794; ketoacyl-synt.
 DR InterPro: IPR001697; Pyruvate_kinase.
 DR Pfam: PF00224; PK; 1.
 DR Pfam: PF02887; PK_C; 1.
 DR PRINTS: PR01050; PYRUVTKINASE.
 DR ProDom: PD001009; Pyruvate_kinase; 1.
 DR TIGRFAMs: TIGR01064; pyruv_kin; 1.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
 DR PROSITE: PS00110; PYRUVATE_KINASE; 1.
 DR GlycoLysis; Kinase; Magnesium; Transferrase.
 KW GlycoLysis; Kinase; Magnesium; Transferrase.
 SQ SEQUENCE 566 AA; 60964 MW; 91852BC0DD4559E CRC64;

Query Match
 Best Local Similarity 89.7%; Score 35; DB 4; Length 566;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8
 Db 258 EOEVRDLR 265

RESULT 14

ID 016715 PRELIMINARY: PRT: 587 AA.

AC 016715: 01-NOV-1996 (TREMBLERL. 01, Created)

DT 01-NOV-1996 (TREMBLERL. 01, Last sequence update)

DT 01-JUN-2002 (TREMBLERL. 21, Last annotation update)

DE Pyruvate kinase (EC 2.7.1.40) (PK) (Fragment).

OS Homo sapiens (Human); Chordata; Craniata; Vertebrata; Euteleostomi;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98182587; PubMed=9522120;

RA Lenzner C., Nurnberg P., Jacobsch G., Thiele B.J.;

RT Complete genomic sequence of the human PK-L/R-gene includes four

RT intragenic polymorphisms defining different haplotype backgrounds of

RT normal and mutant PK-genes.*;

RL DNA Seq. 8:45-53(1997).

CC -1- CATALYTIC ACTIVITY: ATP + PYRUVATE -> ADP + PHOSPHOENOLPYRUVATE.

CC -1- COFACTOR: REQUIRES MAGNESIUM AND POTASSIUM (BY SIMILARITY).

CC -1- PATHWAY: FINAL STEP IN GLYCOLYSIS.

CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE PYRUVATE KINASE FAMILY.

DR EMBL: U67654; AA02535.1; -.

DR HSSP: P11974; IAOE.

DR InterPro: IPR00794; Ketoacyl-synt.

DR InterPro: IPR001697; Pyruvate_kinase.

DR Pfam: PF00224; PK; 1.

DR Pfam: PF02887; PK; 1.

DR PRINTS: PR01050; PYRUVTKINASE.

DR PRODOM: PD001009; Pyruvate_kinase; 1.

DR TIGRFAMs: TIGR01064; pyruv_kin; 1.

DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.

DR PROSITE: PS00110; PYRUVATE_KINASE; 1.

KW Glycolysis; Kinase; Magnesium; transferase.

KW NON_TER

FT SEQUENCE 587 AA; 63260 MW; E6BC7FC7A4B6069 CRC64;

SQ

Query Match 89.7%; Score 35; DB 4; Length 587;

Best Local Similarity 75.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8

Db 279 EODVDRDLR 286

RESULT 15

ID 075758 PRELIMINARY: PRT: 599 AA.

AC 075758: 01-NOV-1998 (TREMBLERL. 08, Created)

DT 01-NOV-1998 (TREMBLERL. 08, Last sequence update)

DT 01-JUN-2002 (TREMBLERL. 21, Last annotation update)

DE Pyruvate kinase (EC 2.7.1.40) (PK).

OS Homo sapiens (Human); Chordata; Craniata; Vertebrata; Euteleostomi;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93075125; PubMed=1445295;

RA Kanno H., Fujii H., Miya S.;

RT Structural analysis of human pyruvate kinase L-gene and

RT identification of the promoter activity in erythroid cells.*;

RT Biochem. Biophys. Res. Commun. 188:516-523(1992).

CC -1- CATALYTIC ACTIVITY: ATP + PYRUVATE -> ADP + PHOSPHOENOLPYRUVATE.

CC -1- COFACTOR: REQUIRES MAGNESIUM AND POTASSIUM (BY SIMILARITY).

CC -1- PATHWAY: FINAL STEP IN GLYCOLYSIS.

CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE PYRUVATE KINASE FAMILY.

DR EMBL: D13243; BAA02515.1; -.

DR EMBL: AB015984; BAA02515.1; JOINED.

DR EMBL: D13233; BAA02515.1; JOINED.

DR EMBL: D13233; BAA02515.1; JOINED.

DR EMBL: D13235; BAA02515.1; JOINED.

DR EMBL: D13236; BAA02515.1; JOINED.

DR EMBL: D13237; BAA02515.1; JOINED.

DR EMBL: D13238; BAA02515.1; JOINED.

DR EMBL: D13239; BAA02515.1; JOINED.

DR EMBL: D13240; BAA02515.1; JOINED.

DR EMBL: D13241; BAA02515.1; JOINED.

DR EMBL: D13242; BAA02515.1; JOINED.

DR HSSP: P11974; IAOE.

DR InterPro: IPR00794; Ketoacyl-synt.

DR InterPro: IPR001697; Pyruvate_kinase.

DR Pfam: PF00224; PK; 1.

DR Pfam: PF02887; PK; 1.

DR PRINTS: PR01050; PYRUVTKINASE.

DR PRODOM: PD001009; Pyruvate_kinase; 1.

DR TIGRFAMs: TIGR01064; pyruv_kin; 1.

DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.

DR PROSITE: PS00110; PYRUVATE_KINASE; 1.

KW Glycolysis; Kinase; Magnesium; transferase.

KW SEQUENCE 599 AA; 64502 MW; 29BCFF812F181A0 CRC64;

SQ

Query Match 89.7%; Score 35; DB 4; Length 599;

Best Local Similarity 75.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EOEIRDLR 8

Db 291 EODVDRDLR 298

Search completed: November 13, 2002, 13:16:12
Job time : 26.7021 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 29.7447 Seconds
(without alignments)
26.879 Million cell updates/sec

Title: US-09-856-086-5

Perfect score: 32

Sequence: 1 KKVHEE 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_101002.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	6	21	Test antigen #6 fo
2	32	100.0	7	21	Human neurofilamen
3	32	100.0	441	19	Human polypeptide,
4	32	100.0	543	22	Human polypeptide,
5	31	96.9	3135	15	P. falciparum tran
6	31	96.9	3135	21	Plasmodium falcipa
7	29	90.6	15	23	Human Orc4lp 47.3
8	29	90.6	76	21	Human secreted pro
9	29	90.6	188	22	Human polypeptide
10	29	90.6	189	22	Novel human secret

11	29	90.6	319	21	AA842280
12	29	90.6	430	23	AB879474
13	29	90.6	465	19	AA554351
14	29	90.6	465	23	AB873394
15	29	90.6	466	21	AA829635
16	29	90.6	466	21	AA829635
17	29	90.6	466	22	AA866348
18	29	90.6	466	22	AA866348
19	29	90.6	475	22	AA895511
20	29	90.6	541	22	AA894625
21	29	90.6	545	22	AA825417
22	29	90.6	598	23	AA848998
23	29	90.6	642	22	AA893669
24	29	90.6	658	22	AA840324
25	29	90.6	846	22	AA85036
26	29	90.6	1390	22	AB862854
27	28	87.5	82	20	AA876517
28	28	87.5	134	22	AB828762
29	28	87.5	134	22	AB819386
30	28	87.5	134	22	AA854712
31	28	87.5	142	18	AA855480
32	28	87.5	147	18	AA855285
33	28	87.5	147	21	AA825305
34	28	87.5	168	23	AB835523
35	28	87.5	189	21	AA855860
36	28	87.5	189	21	AA855965
37	28	87.5	200	21	AA855964
38	28	87.5	201	21	AA855859
39	28	87.5	216	22	AA854786
40	28	87.5	222	20	AA806418
41	28	87.5	248	22	AA818351
42	28	87.5	273	22	AA892564
43	28	87.5	273	22	AA868881
44	28	87.5	358	23	AB855328
45	28	87.5	398	20	AA829335

ALIGNMENTS

RESULT 1	
AA805930	AA805930 standard; peptide: 6 AA.
AC	AA805930:
XX	17-OCT-2000 (first entry)
XX	Test antigen #6 for spongiform and demyelinating disease diagnosis.
DE	Human; cow; myelin; myelin neurofilament; immunogen; antigen;
XX	bovine spongiform encephalopathy; BSE; multiple sclerosis;
KW	Creutzfeldt-Jacob disease; CJD; demyelinating disease; diagnostic test.
XX	
OS	Bos taurus.
OS	Homo sapiens.
PN	WO200031545-A1.
XX	
PD	02-JUN-2000.
XX	
XX	25-NOV-1999; 99WO-GB03936.
XX	
XX	26-NOV-1998; 98GB-0025948.
XX	
PA	(UNLO) KING'S COLLEGE.
XX	
PI	Edlinger A;
XX	
XX	WPI; 2000-400194/34.
DR	
XX	Diagnosing spongiform or demyelinating disease in vertebrates such as
PT	bovine spongiform encephalopathy and Creutzfeldt-Jacob disease comprises

PT assaying a biological sample for myelin and/or myelin neurofilament
 PR antibodies -
 XX
 PS Claim 5; Page 2; 16pp; English.
 CC
 CC The present peptide may be used as a test antigen in a kit for diagnosing
 CC spongiform or demyelinating disease in vertebrates, including bovine
 CC spongiform encephalopathy (BSE); multiple sclerosis (MS) and
 CC Creutzfeld-Jacob disease (CJD). Peptides used in the kit are antigenic
 CC components of myelin or myelin neurofilaments. Biological samples
 CC are assayed for antibodies, especially IGA antibodies, which bind to
 CC myelin and/or myelin neurofilaments or an antigenic peptide, such as the
 CC present sequence. Any reading in excess of two standard deviations of
 CC the healthy controls would indicate a positive response.
 CC
 SQ Sequence 6 AA;
 Query Match 100.0%; Score 32; DB 21; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KKVHEE 6
 DB 1 KKVHEE 6
 RESULT 2
 ID AAB05926
 ID AAB05926 standard; peptide; 7 AA.
 XX
 AC AAB05926;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Test antigen #2 for spongiform and demyelinating disease diagnosis.
 XX
 KM Human; cow; myelin; myelin neurofilament; immunogen; antigen;
 KM bovine spongiform encephalopathy; BSE; multiple sclerosis;
 KM Creutzfeld-Jacob disease; CJD; demyelinating disease; diagnostic test.
 XX
 OS Bos taurus.
 OS Homo sapiens.
 XX
 PN WO200031545-A1.
 XX
 PD 02-JUN-2000.
 XX
 PF 25-NOV-1999; 99WO-GB03936.
 XX
 PR 26-NOV-1998; 98GB-0025948.
 XX
 PA (UNLO) KING'S COLLEGE.
 XX
 PI Ebringer A;
 XX
 DR WPI: 2000-400194/34.
 XX
 PT Diagnosing spongiform or demyelinating disease in vertebrates such as
 PT bovine spongiform encephalopathy and Creutzfeld-Jacob disease comprises
 PT assaying a biological sample for myelin and/or myelin neurofilament
 PT antibodies -
 XX
 PS Claim 5; Page 2; 16pp; English.
 CC
 CC The present peptide may be used as a test antigen in a kit for diagnosing
 CC spongiform or demyelinating disease in vertebrates, including bovine
 CC spongiform encephalopathy (BSE); multiple sclerosis (MS) and
 CC Creutzfeld-Jacob disease (CJD). Peptides used in the kit are antigenic
 CC components of myelin or myelin neurofilaments. Biological samples
 CC are assayed for antibodies, especially IGA antibodies, which bind to
 CC myelin and/or myelin neurofilaments or an antigenic peptide, such as the
 CC present sequence. Any reading in excess of two standard deviations of
 CC the healthy controls would indicate a positive response.

XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 32; DB 21; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KKVHEE 6
 DB 2 KKVHEE 7
 RESULT 3
 ID AAY20612
 ID AAY20612 standard; Protein; 441 AA.
 XX
 AC AAY20612;
 XX
 DT 22-JUL-1999 (first entry)
 XX
 DE Human neurofilament-L wild type protein fragment 2.
 XX
 KM Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KM frameshift mutation; age-related disease; neurodegenerative disorder;
 KM Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KM Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KM diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KM ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KM neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KM glial fibrillary acidic protein; GRAP; p53; semaphorin III; HOPF-1;
 KM bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGp-C; NSP-A;
 KM high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIKSDUNIT UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 DR WPI: 1998-609901/51.
 XX
 PR N-PSDB; AAX75758.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 7; 258pp; English.
 CC
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including the
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule

CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumor antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HSP-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 441 AA;

Query Match 100.0%; Score 32; DB 19; Length 441;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 |||||
 Db 263 KKVHEE 268

RESULT 4

AAW93466
 ID AAW93466 standard; Protein: 543 AA.

XX AAW93466;

XX 06-NOV-2001 (first entry)

XX Human polypeptide, SEQ ID NO: 3132.

XX Human; full length cDNA; cDNA synthesis; oligo-capping.

XX Homo sapiens.

XX EPI130094-A2.

XX 05-SEP-2001.

XX 07-JUL-2000; 2000EP-0114089.

XX 08-JUL-1999; 99JP-0194486.

XX 11-JAN-2000; 2000JP-0118774.

XX 02-MAY-2000; 2000JP-0183765.

XX (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX WPI: 2001-524255/58.

XX N-PSDB: AAK94387.

XX 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -

XX Claim 8; SEQ ID NO 3132; 1380bp + sequence listing; English.

CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 543 AA;

Query Match 100.0%; Score 32; DB 22; Length 543;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 |||||
 Db 224 KKVHEE 229

RESULT 5

AAW57474
 ID AAW57474 standard; Protein: 3135 AA.

XX AAW57474;

XX 20-FEB-1995 (first entry)

XX P. falciparum transmission blocking target antigen Pf230.

XX Protozoan; transmission blocking target antigen; Pf230; malaria;
 KW vaccine.

XX Plasmodium falciparum.

XX WO9417187-A.

XX 04-AUG-1994.

XX 18-JAN-1994; 94WO-US00547.

XX 29-JAN-1993; 93US-0010409.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (USSH) US SEC DEPT HEALTH.

XX Kaslow DC, Williamson KC;

XX WPI: 1994-264101/32.

XX N-PSDB: AAG67190.

XX New Plasmodium falciparum transmission blocking target antigen -
 PT useful in antimalarial vaccines, also related DNA, expression
 PT vectors and transformed cells

XX Claim 6; Page 24; 63pp; English.

XX Pf230 protein is administered to humans to prevent transmission of
 CC malaria by inducing a transmission blocking immune response. It
 CC can also be used to raise antibodies and for T and B cell epitope
 CC mapping. Pf230 induces a high and long-lasting antibody titer and
 CC can be produced in large amounts at low cost.

XX Sequence 3135 AA;

Query Match 96.9%; Score 31; DB 15; Length 3135;
 Best Local Similarity 83.3%; Pred. No. 2.1e+03;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 |||||
 Db 2556 KKVHEE 2561

RESULT 6

AAW18223
 ID AAW18223 standard; Protein: 3135 AA.

XX AAW18223;

XX 07-NOV-2000 (first entry)

XX Plasmodium falciparum chromosome 2 related protein SEQ ID NO:80.

XX Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;
 KW antimalarial; malaria; protozoacide; infection; insecticide.

XX Plasmodium falciparum.

XX MO200025728-A2.
XX
XX PD 11-MAY-2000.
XX
XX PF 05-NOV-1999; 99WO-US26796.
XX PR 05-NOV-1998; 98US-0107131.
XX PA (HOFF/) HOFFMAN S.
XX PA (GARU/) CARUCCI D.
XX PA (GARD/) GARDNER M.
XX PA (VENT/) VENTER J C.
XX
XX PI Hoffman S, Carucci D, Gardner M, Venter JC;
DR WPI: 2000-365347/31.
XX
XX PT Proteins encoded by chromosome 2 of the human malarial parasite,
PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P.falciparum infection -
XX
XX PS Disclosure: Page 192-200; 577pp; English.

The present invention describes proteins and their fragments (I) encoded by chromosome 2 of the human malarial parasite, Plasmodium falciparum. Also described are: (1) nucleotide sequences (II) encoding (I); and (2) vaccines against P. falciparum infection comprising (I) or (II). (I) and (II) are useful for the development of vaccines against P. falciparum infection. (I) and polyclonal antisera or monoclonal antibody raised to immunogens comprising the sequences of (I), are useful in the detection of infection with P. falciparum. Furthermore, (I) (especially when they are rifins or secreted or membrane proteins) can aid the identification of drugs to treat or prevent P. falciparum infection, or they can be used to identify drug resistance in P. falciparum. Sequencing of the plasmodium chromosome 2 and the subsequent identification of proteins encoded by it will help to expand our understanding of parasitic biology, a process hampered by the complexity of the parasitic lifecycle, and provide new targets for vaccine and drug development. Parasite resistance to drugs and mosquito resistance to insecticides have led to a resurgence of malaria in many parts of the world, and there is a pressing need for vaccines and new drugs. AAM70078 to AAM70287 and AAB18144 to AAB18352 represent nucleotide and protein sequences given in the present invention, but which are not specifically mentioned within the specification.

Sequence 3135 AA;

Query Match 96.9%; Score 31; DB 21; Length 3135;
Best Local Similarity 83.3%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0

OY 1 KRYHEE 6
||:||||
DB 2556 KRHHEE 2561

RESULT 7
ABB79475
ID ABB79475 standard; Peptide: 15 AA.
XX
XX AC ABB79475;
XX DT 23-SEP-2002 (first entry)
XX DE Human Orc4Lp 47.3 N-terminal peptide fragment.
XX KW Replication start codon initiation recognition compound; Orc4Lp;
KW HSORC4L; cancer; cytostatic; HIV infection; anti-HIV; virucide;
human; gene therapy.
XX OS Homo sapiens.
XX

FN CNI331153-A.
 PD 16-JAN-2002.
 PF 26-JUN-2000; 2000CN-0116778.
 PR 26-JUN-2000; 2000CN-0116778.
 PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 PI Mao Y, Xie Y;
 DR WPI: 2002-352937/39.
 PT Polypeptide-human replication start codon initiation recognition
 compound subunit Orc4lp (HSORC4L)47.3 and polynucleotide for coding it
 -
 PS Example 5; Page 20 (Disclosure); 33pp; Chinese.
 CC The present invention relates to novel human replication start
 CC codon initiation recognition compound subunit Orc4lp (HSORC4L) 47.3
 CC (see ABB74474). This protein and its coding sequence are useful in
 CC the treatment of diseases such as cancer and HIV infection. The
 CC present sequence is an N-terminal peptide fragment of the protein,
 CC which was used in an example from the invention.
 SQ Sequence 15 AA;
 Query Match 90.6%; Score 29; DB 23; Length 15;
 Best Local Similarity 83.3%; Pred. No. 23;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0
 QY 1 KKVHEE 6
 ||:||||
 Db 10 KKHHEE 15
 RESULT 8
 AAG03533
 ID AAG03533 standard; Protein; 76 AA.
 XX
 AC AAG03533;
 DT 06-OCT-2000 (first entry)
 DE Human secreted protein, SEQ ID NO: 7614.
 DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 PD 06-SEP-2000.
 PF 21-FEB-2000; 2000EP-0200610.
 PR 26-FEB-1999; 99US-0122487.
 PA (GEST) GENSET.
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 DR WPI: 2000-500381/45.
 DR N-PSDB: AAC03539.
 PS New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 diagnostic, forensic, gene therapy and chromosome mapping procedures -
 Claim 13; SEQ ID 7614; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used in diagnostic, forensic, gene therapy and
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.

SO Sequence 76 AA:

Query Match 90.6%; Score 29; DB 21; Length 76;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 ||:||||
 DB 62 KKLHEE 67

RESULT 9

AA042110
 ID AA042110 standard; Protein; 188 AA.

AC AA042110;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 7041.

XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
 XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 XX leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Maundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX Zhao QH, Zhou P, Goodfich R, Drmanac RT;

XX MPI: 2001-442253/47.

XX N-PSDB; AA161266.

XX Novel nucleic acids and polypeptides, useful for treating disorders

XX PT such as central nervous system injuries -

XX Example 2; SEQ ID NO 7041; 10078bp; English.

XX The invention relates to human nucleic acids (AA15798-AA161369) and
 CC the encoded polypeptides (AA038642-AA042113) with neotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.

SO Sequence 188 AA:

Query Match 90.6%; Score 29; DB 22; Length 188;
 Best Local Similarity 83.3%; Pred. No. 3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKVHEE 6
 ||:||||
 DB 106 KKLHEE 111

RESULT 10

AA029773
 ID AA029773 standard; Protein; 189 AA.

AC AA029773;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #264.

XX Human; vaccination; gene therapy; nutritional supplement;
 XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN WO200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US08656.

PR 18-APR-2000; 2000US-0552929.

PR 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX MPI: 2001-611725/70.

XX Nucleic acids encoding a range of human polypeptides, useful in genetic

XX vaccination, testing and therapy -

XX Claim 20; Page 192; 765bp; English.

XX The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising

CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate hematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukemias. AA029510-AA03304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.

XX Sequence 189 AA:

Query Match 90.6%; Score 29; DB 22; Length 189;
Best Local Similarity 83.3%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKHHEE 6
111111
Db 106 KKLHEE 111

RESULT 11

AA042280 standard; Protein: 319 AA.

XX AAB42280;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2044 polypeptide sequence SEQ ID NO:4088.

XX Human; open reading frame; ORFX; detection: cytostatic; hepatotropic;
XX vulnery; antiparkinsonian; antiparkinsonian; neurotrophic; neuroprotective;
XX anticonvulsant; osteopathic; antiallergic; immunosuppressant; cardiant;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatological; immunosuppressive; antitubercular;
XX antiviral; antibacterial; antifungal; antineoplastic; antihypertensive;
XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive.

XX Homo sapiens.

XX WO200058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US08621.

XX 31-MAR-1999; 99US-0127607.

XX 02-APR-1999; 99US-0127636.

XX 05-APR-1999; 99US-0127728.

XX 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shinkens RA, Leach M;

XX WPI: 2000-602362/57.

XX N-PSDB: AAC76489.

XX Novel nucleic acids and peptides derived from open reading frame X,
XX useful for treating e.g. cancers, proliferative disorders,
XX neurodegenerative disorders and cardiovascular disease -
XX Claim 11; Page 3278; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
XX CC

CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnery;
CC antiparkinsonian; neurotrophic; neuroprotective;
CC osteopathic; anticonvulsant; antiallergic; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antineoplastic; antibacterial; antiviral; antifungal; antirheumatic;
CC antihypertensive; antianemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC the pathological conditions associated with an ORFX-associated disorder. The
CC nucleic acids can be used to express ORFX proteins in gene therapy
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX Sequence 319 AA:

Query Match 90.6%; Score 29; DB 21; Length 319;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKHHEE 6
111111
Db 152 KKLHEE 157

RESULT 12

ABB79474 standard; Protein: 430 AA.

XX ABB79474;

DT 23-SEP-2002 (first entry)

XX Human Orc4Lp 47.3 protein.

XX Replication start codon initiation recognition compound; Orc4Lp;
XX HsORC4L; cancer; cytostatic; HIV infection; anti-HIV; virucide;
XX human; gene therapy.

XX Homo sapiens.

XX CN1331153-A.

XX 16-JAN-2002.

XX 26-JUN-2000; 2000CN-0116778.

XX 26-JUN-2000; 2000CN-0116778.

XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX Mao Y, Xie Y;

XX WPI: 2002-352937/39.

XX N-PSDB: ABB84122.

XX Polypeptide-human replication start codon initiation recognition
XX compound subunit Orc4Lp (HsORC4L)47.3 and polynucleotide for coding it
XX PT

XX Claim 1; Page 28-29 (Disclosure); 33pp; Chinese.

XX The present sequence is the protein sequence of novel human
XX replication start codon initiation recognition compound subunit
XX Orc4Lp (HsORC4L) 47.3. This protein and its coding sequence are
XX useful in the treatment of diseases such as cancer and HIV
XX CC

CC Infection.
 XX
 SQ Sequence 430 AA;
 Query Match 90.6%; Score 29; DB 23; Length 430;
 Best Local Similarity 83.3%; Pred. No. 6.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KKYHEE 6
 11:1111
 Db 10 KKLHEE 15
 RESULT 13
 AAW54351
 ID AAW54351 standard; protein; 465 AA.
 AC AAW54351;
 XX
 DT 14-AUG-1998 (first entry)
 XX
 DE Vimentin.
 KW Endometrium; hyperplasia; adenocarcinoma; proliferative phase;
 KM 2D gel electrophoresis; detection.
 XX
 OS Homo sapiens.
 PN W09810291-A1.
 PD 12-MAR-1998.
 XX
 PF 05-SEP-1997; 97MO-G802394.
 XX
 PR 08-APR-1997; 97GB-0007132.
 PR 06-SEP-1996; 96GB-0018600.
 PA (CLIN-) CENT CLINICAL & BASIC RES.
 PI Byrjalsen I, Fey SJ, Larsen P;
 DR WPI: 1998-207057/18.
 PT Biochemical markers of human endometrium - useful for, e.g.
 diagnosis of hyperplasia and adenocarcinoma
 PS Disclosure: Page 20; 77pp; English.
 CC Proteins AAW54349-W54364 are examples of proteins produced in the
 endometrium during the hyperplasia, adenocarcinoma or proliferative
 phase of the endometrium. The presence and quantities of these proteins
 can be detected using 2D gel electrophoresis comparison of cell lysates.
 CC The proteins can be used as biochemical markers to detect the phase of
 the endometrium and can be measured in body fluids, obviating the need
 for endometrial biopsies.
 SQ Sequence 465 AA;
 Query Match 90.6%; Score 29; DB 19; Length 465;
 Best Local Similarity 83.3%; Pred. No. 7.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KKYHEE 6
 11:1111
 Db 234 KKLHEE 239
 RESULT 14
 ABB77394
 ID ABB77394 standard; protein; 465 AA.
 AC ABB77394;
 XX

DT 11-JUL-2002 (first entry)
 XX
 DE Human vimentin.
 KM Human; dermatological; skin stress; aging; spondin 2; cathepsin L;
 KM actin gamma 1; vimentin; fibroblast; skin; cosmetic; pharmaceutical.
 OS Homo sapiens.
 PN DE10050274-A1.
 PD 18-APR-2002.
 XX
 PF 09-OCT-2000; 2000DE-1050274.
 XX
 PR 09-OCT-2000; 2000DE-1050274.
 XX
 PA (HENK) HENKEL KGAA.
 PI Petersohn D, Schmitt G, Foerster T;
 DR WPI: 2002-373046/41.
 XX
 PT In vitro assays for skin stress and skin aging includes determination
 of spondin 2, cathepsin L, actin gamma 1 and vimentin fragments
 secreted by skin fibroblasts
 PS Claim 5; Page 12; 14pp; German.
 CC The invention relates to in vitro methods for the detection of skin
 stress and/or skin aging in humans and animals based on the
 determination of spondin 2, cathepsin L, actin gamma 1 or vimentin
 fragments secreted by fibroblast from the skin under test. Use of the
 methods in a test for potential cosmetics and pharmaceuticals with an
 effect on these skin conditions and products containing vimentin
 fragments are also included. Products containing vimentin fragments
 are effective in the regulation, especially maintenance, of skin
 homeostasis.
 SQ Sequence 465 AA;
 Query Match 90.6%; Score 29; DB 23; Length 465;
 Best Local Similarity 83.3%; Pred. No. 7.4e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KKYHEE 6
 11:1111
 Db 235 KKLHEE 240
 RESULT 15
 AAB29635
 ID AAB29635 standard; protein; 466 AA.
 AC AAB29635;
 XX
 DT 21-FEB-2001 (first entry)
 DE Human pollinosis-associated gene 795-encoded protein, SEQ ID NO:26.
 KW Human; pollinosis-associated gene 795; vimentin homologue;
 KM IgE; immunoglobulin E; cedar pollen allergy; T-cell; reduced expression;
 detection; diagnosis; drug screening; allergic disease.
 OS Homo sapiens.
 PN W0200065050-A1.
 PD 02-NOV-2000.
 XX
 PF 26-APR-2000; 2000WO-JP02734.
 PR 27-APR-1999; 99JP-0120494.

XX (GENO-) GENOX RES INC.
PA (EISA) EISAI CO LTD.

XX
PI Nagasu T, Sugita Y, Kashiwabara T, Oshida T, Obayashi M, Gunji S;
PI Obayashi I, Imai Y, Yoshida N, Ogawa K, Matsui K, Takahashi E;
PI Yokoi A;

XX
DR WPI: 2000-687343/67.
DR N-PSDB: AAC64226.

XX
PT Pollinosis-associated gene 795 undergoing significantly low expression
PT in subjects with high cedar pollen-specific IgE levels, useful in
PT diagnosis of allergic diseases and screening drug candidates -

PS Page 64-67; Claim 13; 73pp; Japanese.

XX
CC The invention relates to the human pollinosis-associated gene 795 which
CC exhibits significantly reduced expression in the T-cells of individuals
CC with high cedar pollen-specific IgE (immunoglobulin E) levels. The gene
CC was isolated from T-cells from individuals allergic to cedar pollen using
CC the differential display method. Pollinosis-associated gene 795 has
CC homology with the human vimentin gene. The invention also relates also
CC relates to the protein encoded by pollinosis gene 795, to expression
CC constructs and host cells comprising pollinosis-associated gene 795
CC nucleic acids; pollinosis-associated gene 795 primers and probes;
CC antibodies against the protein encoded by the gene; methods of detection
CC of pollinosis-associated gene 795 nucleic acids; and a method of
CC diagnosis of allergic diseases via the detection of pollinosis-associated
CC gene 795 nucleic acids. The invention additionally encompasses methods of
CC screening drug candidates for the treatment of allergic disease by
CC measuring the expression of pollinosis-associated gene 795 in pollen
CC antigen-stimulated T-cells in the presence of a test compound relative to
CC a control. Pollinosis-associated gene 795 is useful in the diagnosis of
CC allergic diseases and in the screening of drug candidates for the
CC treatment of such diseases. The present sequence represents a
CC protein encoded by human pollinosis-associated gene 795.

XX
SQ Sequence 466 AA:

Query Match 90.6%; Score 29; DB 21; Length 466;

Best Local Similarity 83.3%; Pred. No. 7.4e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KKYHHE 6

Db 235 KKLHHE 240

Search completed: November 13, 2002, 13:22:10
Job time : 30.7447 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 29.7447 Seconds
(without alignments)
26.879 Million cell updates/sec

Title: US-09-856-086-1

Perfect score: 29

Sequence: 1 NEALEX 6

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	29	100.0	6	Test antigen #1 fo
2	29	100.0	103	Human secreted pro
3	29	100.0	183	Serratia marcescen
4	29	100.0	220	Arabidopsis thalia
5	29	100.0	272	Arabidopsis thalia
6	29	100.0	290	Arabidopsis thalia
7	29	100.0	290	Herbicideally activ
8	29	100.0	323	Arabidopsis thalia
9	29	100.0	413	Lactococcus lactis
10	29	100.0	441	Human neurofilamen

11	29	100.0	543	22	AA093466
12	29	100.0	550	20	AAV34552
13	29	100.0	599	20	AAV34551
14	29	100.0	614	20	AAV34417
15	29	100.0	1151	22	AB061598
16	29	100.0	1829	18	AAW29322
17	27	93.1	94	22	ABB23413
18	27	93.1	94	22	ABB22990
19	27	93.1	130	20	AAV35002
20	27	93.1	144	22	AAU34141
21	27	93.1	147	22	AAU36805
22	27	93.1	147	22	AAU37375
23	27	93.1	147	22	AAU37561
24	27	93.1	191	21	AAU14801
25	27	93.1	253	22	ABB62923
26	27	93.1	326	21	AAU14800
27	27	93.1	336	21	AAU14799
28	27	93.1	586	22	AAU95313
29	27	93.1	620	22	ABB91254
30	27	93.1	770	22	AAU93560
31	27	93.1	886	9	AAU80345
32	27	93.1	886	23	ABP35643
33	26	89.7	18	22	ABB30757
34	26	89.7	18	22	ABB35936
35	26	89.7	18	22	ABB21343
36	26	89.7	18	22	AAU67737
37	26	89.7	18	22	AAU69117
38	26	89.7	18	22	AAU16949
39	26	89.7	18	22	AAU29436
40	26	89.7	18	22	AAU04646
41	26	89.7	18	22	ABG38713
42	26	89.7	30	23	AAU16140
43	26	89.7	97	22	ABG21567
44	26	89.7	107	22	AAU14183
45	26	89.7	118	21	AAU00936

ALIGNMENTS

RESULT 1
ID AAB05925 standard; peptide; 6 AA.
AC AAB05925;
XX 17-OCT-2000 (first entry)
DT
XX Test antigen #1 for spongiform and demyelinating disease diagnosis.
DE
XX Human; cow; myelin; myelin neurofilament; immunogen; antigen;
KW bovine spongiform encephalopathy; BSE; multiple sclerosis;
KW Creutzfeld-Jacob disease; CJD; demyelinating disease; diagnostic test.
XX
OS Bos taurus.
OS Homo sapiens.
OS WO200031545-A1.
XX
XX 02-JUN-2000.
XX
XX 25-NOV-1999; 99WO-GB03936.
XX
XX 26-NOV-1998; 98GB-0025948.
XX
XX (UNLO) KING'S COLLEGE.
XX
XX Ebringer A;
XX
XX WPI; 2000-400194/34.
XX
XX Diagnosing spongiform or demyelinating disease in vertebrates such as bovine spongiform encephalopathy and Creutzfeld-Jacob disease comprises

PT assaying a biological sample for myelin and/or myelin neurofilament
 PT antibodies -
 XX
 PS Claim 5; Page 2; 16pp; English.
 CC The present peptide may be used as a test antigen in a kit for diagnosing
 CC spongiform or demyelinating disease in vertebrates, including bovine
 CC spongiform encephalopathy (BSE), multiple sclerosis (MS) and
 CC Creutzfeldt-Jacob disease (CJD). Peptides used in the kit are antigenic
 CC components of myelin or myelin neurofilaments. Biological samples
 CC are assayed for antibodies, especially IGA antibodies, which bind to
 CC myelin and/or myelin neurofilaments or an antigenic peptide, such as the
 CC present sequence. Any reading in excess of two standard deviations of
 CC the healthy controls would indicate a positive response.
 XX
 SQ Sequence 6 AA;
 Query Match 100.0%; Score 29; DB 21; Length 6;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NEALEK 6
 |||||
 Db 1 NEALEK 6
 RESULT 2
 AAG02130
 ID AAG02130 standard; Protein; 103 AA.
 XX
 AC AAG02130;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 6211.
 XX
 KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 DR N-PSDB; AAC02136.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PT
 PS Claim 13; SEQ ID 6211; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and

CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 103 AA;
 Query Match 100.0%; Score 29; DB 21; Length 103;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 NEALEK 6
 |||||
 Db 80 NEALEK 85
 RESULT 3
 AAY50040
 ID AAY50040 standard; protein; 183 AA.
 XX
 AC AAY50040;
 XX
 DT 19-JAN-2000 (first entry)
 XX
 DE Serratia marcescens translation initiation factor IF3.
 XX
 KM Translation initiation factor; IF3; bacterial; screening;
 KM antibacterial; antibiotic; drug target; initiation codon; recognition;
 KM ubiquitous; conserved; broad spectrum; infection; contamination;
 KM sterilisation.
 XX
 OS Serratia marcescens.
 XX
 PN WO952357-A1.
 PD 21-OCT-1999.
 XX
 PF 14-APR-1999; 99WO-US08134.
 XX
 PR 14-APR-1998; 98US-0081736.
 XX
 PA (RIBO-) RIBOGENE INC.
 XX
 PI Dammel CS, Watson JC, Hernandez VT;
 XX
 DR WPI; 1999-620242/53.
 XX
 PT Assays for inhibitors of bacterial translation initiation factor 3 -
 PT
 PS Disclosure; Fig 1; 66pp; English.
 XX
 CC This sequence represents Serratia marcescens translation initiation
 CC factor IF3. IF3 is an essential factor, acting to prevent association
 CC of ribosomal subunits and to recognise the correct initiation codon.
 CC IF3 can be used as a drug target for screening for test compounds that
 CC inhibit the activity of IF3 which comprise using a reporter gene
 CC system in whole cells and detecting the ability of IF3 to discriminate
 CC against translation initiation at an atypical start codon of the
 CC reporter gene. IF3 is apparently ubiquitous and conserved throughout
 CC the bacterial kingdom. It is therefore likely that any compounds which
 CC are effective at inhibiting IF3 of one bacterial species will have an
 CC inhibitory effect on the IF3 or its functional equivalents in a wide
 CC range of bacteria. Importantly, IF3 has no functional homologue in
 CC mammalian cells. As a result, the potential toxicity of IF3 inhibitors
 CC is likely to be low. Compounds identified by the method that inhibit
 CC bacterial IF3 are useful for inhibiting the growth of bacteria,
 CC especially for treating an infectious disease in a human or a companion
 CC or livestock animal. The compound may be adjunctionally administered with
 CC a second antibacterial compound. IF3 inhibitors identified by the
 CC method are also useful for sterilising bacteria-contaminated objects
 CC or material. The compounds identified can penetrate and specifically
 CC kill the pathogenic bacterial cell, or arrest its growth without also
 CC adversely affecting its human, animal or plant host. This is because
 CC the target is highly selective as no homologous mammalian counterpart

CC exists. The high-throughput primary screen allows for the easy visible
CC identification of positive hits.
XX
SQ Sequence 183 AA;
Query Match 100.0%; Score 29; DB 20; Length 183;
Best Local Similarity 100.0%; Pred.No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NEALEX 6
DB 40 NEALEX 45
RESULT 4
AAG44373
ID AAG44373 standard; Protein; 220 AA.
XX
AC AAG44373;
XX
DE 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55574.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137526.

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145818.
PR 27-JUL-1999; 99US-0145819.
PR 28-JUL-1999; 99US-0145951.
PR 28-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.

```

PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151033.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155133.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157763.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159299.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match      100.0%; Score 29; DB 21; Length 220;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NEALEX 6
Db 53 NEALEX 58

RESULT 5
ABBG1406
ID ABBG1406 standard; Protein; 272 AA.
XX
AC ABBG1406;
XX
DI 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 11010.
XX
KW Drosophila: developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR MPI. 2001-656860/75.
XX
DR N-PSDB; ABL05509.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 11010; 21pp + Sequence Listing; English.
XX
PS
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABL30511), expressed DNA
CC sequences (ABH57737-ABH72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 272 AA;

Query Match      100.0%; Score 29; DB 22; Length 272;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NEALEX 6
Db 48 NEALEX 53

RESULT 6
AAG44372
ID AAG44372 standard; Protein; 290 AA.
XX
AC AAG44372;
XX

```

DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55573.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
PN
XX
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128334.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141827.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145324.
PR 26-JUL-1999; 99US-0145376.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.


```

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157573.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158365.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161923.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match          100.0%; Score 29; DB 21; Length 290;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 NEALEX 6
DB 123 NEALEX 128

```

```

RESULT 7
ABB91341
ID ABB91341 standard; Protein; 290 AA.
XX ABB91341;
XX 31-MAY-2002 (first entry)
XX Herbicidally active polypeptide SEQ ID NO 552.
DE Herbicidally active polypeptide SEQ ID NO 552.
XX Herbicidally active polypeptide SEQ ID NO 552.
XX Herbicidally active polypeptide SEQ ID NO 552.
XX Herbicidally active polypeptide SEQ ID NO 552.
OS Arabidopsis thaliana.

```

```

XX WO200210210-A2.
XX 07-FEB-2002.
XX 28-AUG-2001; 2001WO-EP09892.
XX 28-AUG-2001; 2001WO-EP09892.
XX (FARB ) BAYER AG.
XX Tietjen K, Weidner M;
XX WPI; 2002-269010/31.
XX Identifying plant target proteins for herbicidally active compounds,
XX comprising aligning and comparing nucleic acid or amino acid sequences
XX from plant with nucleic acid or amino acid sequences from non plant
XX organisms -
XX Claim 5; SEQ ID NO 552; 261pp + Sequence Listing; English.
XX The invention relates to identifying target proteins
XX (ABB90790-ABB94016) for herbicidally active compounds, comprising
XX aligning and comparing nucleic acid or amino acid sequences from plant
XX with nucleic acid or amino acid sequences from non-plant organisms using
XX suitable search parameters, where plant sequences having an E-value
XX greater by a factor of 3 than the E-value of most similar non-plant
XX sequences are selected. The polypeptides or nucleic acids encoding them
XX are useful for identifying modulators. The identified modulators are
XX useful as herbicides.
SQ Sequence 290 AA;

```

```

Query Match          100.0%; Score 29; DB 23; Length 290;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 NEALEX 6
DB 123 NEALEX 128

```

```

RESULT 8
AAG44371
ID AAG44371 standard; Protein; 323 AA.
XX AAG44371;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55572.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55572.
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.

```

PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0133484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0138119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145115.
PR 23-JUL-1999; 99US-0145148.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151330.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

PR 13-OCT-1999; 99US-0159295;
 PR 14-OCT-1999; 99US-0159329;
 PR 14-OCT-1999; 99US-0159330;
 PR 14-OCT-1999; 99US-0159331;
 PR 14-OCT-1999; 99US-0159337;
 PR 14-OCT-1999; 99US-0159638;
 PR 18-OCT-1999; 99US-0159584;
 PR 21-OCT-1999; 99US-0160741;
 PR 21-OCT-1999; 99US-0160767;
 PR 21-OCT-1999; 99US-0160768;
 PR 21-OCT-1999; 99US-0160770;
 PR 21-OCT-1999; 99US-0160814;
 PR 21-OCT-1999; 99US-0160815;
 PR 22-OCT-1999; 99US-0160980;
 PR 22-OCT-1999; 99US-0160981;
 PR 22-OCT-1999; 99US-0160989;
 PR 25-OCT-1999; 99US-0161404;
 PR 25-OCT-1999; 99US-0161405;
 PR 25-OCT-1999; 99US-0161406;
 PR 26-OCT-1999; 99US-0161359;
 PR 26-OCT-1999; 99US-0161360;
 PR 26-OCT-1999; 99US-0161361;
 PR 28-OCT-1999; 99US-0161920;
 PR 28-OCT-1999; 99US-0161993;
 PR 28-OCT-1999; 99US-0161993;
 PR 29-OCT-1999; 99US-0162142.

Query Match Best Local Similarity 100.0%; Score 29; DB 21; Length 323;
 Pred. No. 2,5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 Db 156 NEALEK 161

RESULT 9

ABB54856 ID ABB54856 standard; Protein; 413 AA.

XX ABB54856;

XX 16-MAY-2002 (first entry)

XX Lactococcus lactis protein purd.

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

XX Lactococcus lactis IL1403.

XX FR2807446-A1.

XX 12-OCT-2001.

XX 11-APR-2000; 2000FR-0004630.

XX 11-APR-2000; 2000FR-0004630.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX Bolotline A, Sorokline A, Renault P, Ehrlich SD;

XX WPI; 2002-043418/06.

XX New nucleotide sequence useful in the identification or Lactococcus

XX lactis and related species -

XX Claim 6; SEQ ID No 1558; 2504pp; French.

XX The present invention is related to a Lactococcus lactis nucleotide
 CC sequence (ABA90521) and related proteins (ABB53300-ABB5621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or

CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

XX Sequence 413 AA;

Query Match Best Local Similarity 100.0%; Score 29; DB 23; Length 413;
 Pred. No. 3,2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 Db 240 NEALEK 245

RESULT 10
 AAY20612 ID AAY20612 standard; Protein; 441 AA.

XX AAY20612;

XX 22-JUL-1999 (first entry)

XX Human neurofilament-L wild type protein fragment 2.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin 1; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HDPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGp-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.

XX Homo sapiens.

XX WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACADEMY OF SCI.

XX (UYRO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX WPI; 1998-609901/51.

XX N-PSDB; AAX75758.

XX Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA

XX Disclosure; Figure 7; 258pp; English.

XX This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II

CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC protein Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HSPF-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.

SO Sequence 441 AA;

Query Match 100.0%; Score 29; DB 19; Length 441;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 385 NEALEK 390

RESULT 11

AA93466 standard; Protein; 543 AA.

AA93466;

06-NOV-2001 (first entry)

Human polypeptide, SEQ ID NO: 3132.

Human; full length cDNA; cDNA synthesis; oligo-capping.

Homo sapiens.

EP130094-A2.

05-SEP-2001.

07-JUL-2000; 2000EP-0114089.

08-JUL-1999; 99JP-0194486.

11-JAN-2000; 2000JP-0118774.

02-MAY-2000; 2000JP-0183765.

(HELT-) HELIX RES INST.

Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

WPI: 2001-524255/58.

N-PSDB; AAK94387.

830 Primers useful for synthesizing full length cDNA clones and their
 use in genetic manipulation -

Claim 8; SEQ ID NO 3132; 1380bp + sequence listing; English.

CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5' - and 3' - ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesized by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 543 AA;

Query Match 100.0%; Score 29; DB 22; Length 543;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 326 NEALEK 331

RESULT 12

AA934552 standard; Protein; 550 AA.

AA934552;

25-AUG-1999 (first entry)

Porphyromonas gingivalis protein PG8.

Porphyromonas gingivalis; PG; periodontal disease; gingivitis;

vacuine; antigenic.

Porphyromonas gingivalis.

WO929870-A1.

17-JUN-1999.

10-DEC-1998; 98WO-AU01023.

04-AUG-1998; 98AU-0005028.

10-DEC-1997; 97AU-0000839.

31-DEC-1997; 97AU-0001182.

30-JAN-1998; 98AU-0001546.

10-MAR-1998; 98AU-0002264.

08-APR-1998; 98AU-0002911.

23-APR-1998; 98AU-0003128.

05-MAY-1998; 98AU-0003358.

WPI: 1999-385613/32.

Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
 Ross BC, Rothel LJ, Webb EA;

N-PSDB; AAK91770.

Claim 1; Page 545-546; 588pp; English.

CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
 CC Porphyromonas gingivalis (Pg) polypeptide sequences given in AAX93318 to
 CC AAX934583. AAX91802 to AAX91989 represent PCR primers used in the
 CC isolation of the Pg polypeptides. The Pg polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The Pg polypeptides can be
 CC used as vaccines especially against Porphyromonas gingivalis. Probes can
 CC be used to detect Porphyromonas gingivalis in standard hybridisation
 CC assays. Porphyromonas gingivalis is involved in periodontal disease
 CC especially gingivitis.

SO Sequence 550 AA;

Query Match 100.0%; Score 29; DB 20; Length 550;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 72 NEALEK 77

RESULT 13

AAV34551 standard; Protein; 599 AA.

AAV34551;

25-AUG-1999 (first entry)

Porphyromonas gingivalis protein PG8.

Porphyromonas gingivalis; PG; periodontal disease; gingivitis;

vacaine; antigenic.

Porphyromonas gingivalis.

WO9929870-A1.

17-JUN-1999.

10-DEC-1998; 98WO-AU01023.

04-AUG-1998; 98AU-0005028.

10-DEC-1997; 97AU-0000839.

31-DEC-1997; 97AU-0001182.

30-JAN-1998; 98AU-0001546.

10-MAR-1998; 98AU-0002264.

09-APR-1998; 98AU-0002911.

23-APR-1998; 98AU-0003128.

05-MAY-1998; 98AU-0003338.

22-MAY-1998; 98AU-0003654.

29-JUL-1998; 98AU-0004917.

(CSLC-) CSL LTD.

Agilus CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;

Rose BC, Rothel LJ, Webb EA;

WPI. 1999-385613/32.

N-PSDB; AAV31769.

Antigenic Porphyromonas gingivalis peptides for preventing

gingivitis

Claim 1; Page 543-544; 588pp; English.

AAV31536 to AAV31801 encode two hundred and sixty six antigenic

Porphyromonas gingivalis (PG) polypeptide sequences given in AAV34318 to

AAV34583. AAV31802 to AAV31989 represent PCR primers used in the

isolation of the PG polypeptides. The PG polypeptides have antibacterial

activity with a vaccine mechanism of action. The PG polypeptides can be

used as vaccines especially against Porphyromonas gingivalis. Probes can

be used to detect Porphyromonas gingivalis in standard hybridisation

assays. Porphyromonas gingivalis is involved in periodontal disease

AAV34417 standard; Protein; 614 AA.

AAV34417;

25-AUG-1999 (first entry)

Porphyromonas gingivalis protein PG8.

Porphyromonas gingivalis; PG; periodontal disease; gingivitis;

vacaine; antigenic.

Porphyromonas gingivalis.

WO9929870-A1.

17-JUN-1999.

10-DEC-1998; 98WO-AU01023.

04-AUG-1998; 98AU-0005028.

10-DEC-1997; 97AU-0000839.

31-DEC-1997; 97AU-0001182.

30-JAN-1998; 98AU-0001546.

10-MAR-1998; 98AU-0002264.

09-APR-1998; 98AU-0002911.

23-APR-1998; 98AU-0003128.

05-MAY-1998; 98AU-0003338.

22-MAY-1998; 98AU-0003654.

29-JUL-1998; 98AU-0004917.

(CSLC-) CSL LTD.

Agilus CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;

Rose BC, Rothel LJ, Webb EA;

WPI. 1999-385613/32.

N-PSDB; AAV31635.

Antigenic Porphyromonas gingivalis peptides for preventing

gingivitis

Claim 1; Page 389-390; 588pp; English.

AAV31536 to AAV31801 encode two hundred and sixty six antigenic

Porphyromonas gingivalis (PG) polypeptide sequences given in AAV34318 to

AAV34583. AAV31802 to AAV31989 represent PCR primers used in the

isolation of the PG polypeptides. The PG polypeptides have antibacterial

activity with a vaccine mechanism of action. The PG polypeptides can be

used as vaccines especially against Porphyromonas gingivalis. Probes can

be used to detect Porphyromonas gingivalis in standard hybridisation

assays. Porphyromonas gingivalis is involved in periodontal disease

especially gingivitis.

Query Match 100.0%; Score 29; DB 20; Length 599;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 121 NEALEK 126

RESULT 14

AAV34417

Query Match 100.0%; Score 29; DB 20; Length 614;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 136 NEALEK 141

RESULT 15

ABB61598

ABB61598 standard; Protein; 1151 AA.

ABB61598;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 11586.

XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX

OS Drosophila melanogaster.
 XX

PN WO200171042-A2.
 XX

PD 27-SEP-2001.
 XX

PF 23-MAR-2001; 2001WO-US09231.
 XX

PR 23-MAR-2000; 2000US-191637P.
 XX

PR 11-JUL-2000; 2000US-0614150.
 XX

PA (PEKE) PE CORP NY.
 XX

PI Venter JC, Adams M, Li PWD, Myers EW;
 XX

DR WPI; 2001-656860/75.
 XX

DR N-PSDB; ABL05701.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX

PS Disclosure: SEQ ID NO 11586; 21bp + Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC

XX
 SQ Sequence 1151 AA;

Query Match 100.0%; Score 29; DB 22; Length 1151;

Best Local Similarity 100.0%; Pred. No. 9.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 |||||
 DB 1099 NEALEK 1104

Search completed: November 13, 2002, 13:22:04
 Job time : 30.7447 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 8.68085 Seconds
(without alignments)
20.336 Million cell updates/sec

Title: US-09-856-086-1

Perfect score: 29
Sequence: 1 NEALEK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/3A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/3B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/3A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/3B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/3C.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/3D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	29	100.0	160 2	US-08-726-306A-183 Sequence 183, App
2	26	89.7	253 3	US-08-792-014-1 Sequence 1, Appl
3	26	89.7	253 4	US-09-443-948-1 Sequence 1, Appl
4	26	89.7	565 1	US-08-961-083-218 Sequence 218, App
5	25	86.2	16 1	US-08-426-627-12 Sequence 12, Appl
6	25	86.2	37 2	US-08-743-200-10 Sequence 10, Appl
7	25	86.2	135 1	US-08-426-627-15 Sequence 15, Appl
8	25	86.2	212 2	US-08-477-396A-2 Sequence 2, Appl
9	25	86.2	252 4	US-09-134-001C-3754 Sequence 3754, Ap
10	25	86.2	314 2	US-08-353-476-71 Sequence 71, Appl
11	25	86.2	399 4	US-09-134-001C-4571 Sequence 4571, Ap
12	25	86.2	429 4	US-09-362-473-12 Sequence 12, Appl
13	25	86.2	529 1	US-08-484-105-8 Sequence 8, Appl
14	25	86.2	539 1	US-08-484-106-8 Sequence 8, Appl
15	25	86.2	683 1	US-07-878-960-2 Sequence 2, Appl
16	25	86.2	683 2	US-08-477-396A-17 Sequence 17, Appl
17	25	86.2	777 2	US-08-477-396A-4 Sequence 4, Appl
18	25	86.2	779 1	US-08-426-627-4 Sequence 4, Appl
19	25	86.2	779 1	US-08-426-627-24 Sequence 24, Appl
20	25	86.2	811 1	US-08-426-627-2 Sequence 2, Appl
21	25	86.2	811 1	US-08-426-627-22 Sequence 22, Appl
22	25	86.2	836 1	US-08-426-627-6 Sequence 6, Appl
23	25	86.2	837 1	US-08-426-627-23 Sequence 23, Appl
24	25	86.2	1285 1	US-07-382-945-2 Sequence 2, Appl
25	25	86.2	1285 1	US-08-453-141-2 Sequence 2, Appl
26	25	86.2	1285 3	US-08-293-314-2 Sequence 2, Appl
27	25	86.2	1356 4	US-09-770-170-6 Sequence 6, Appl

28	24	82.8	9	2	US-08-372-100C-5	Sequence 5, Appl
29	24	82.8	9	3	US-08-159-339A-611	Sequence 611, App
30	24	82.8	9	4	US-08-865-511B-5	Sequence 5, Appl
31	24	82.8	9	4	US-08-865-511B-19	Sequence 19, Appl
32	24	82.8	9	5	PCR-US95-17085-5	Sequence 5, Appl
33	24	82.8	10	1	US-08-787-547-65	Sequence 65, Appl
34	24	82.8	22	3	US-08-940-095-94	Sequence 94, Appl
35	24	82.8	22	3	US-08-940-093-94	Sequence 94, Appl
36	24	82.8	22	3	US-08-940-096-94	Sequence 94, Appl
37	24	82.8	22	4	US-09-465-719-94	Sequence 94, Appl
38	24	82.8	22	4	US-09-453-605-94	Sequence 94, Appl
39	24	82.8	22	4	US-09-453-838-94	Sequence 94, Appl
40	24	82.8	26	2	US-08-690-011A-37	Sequence 37, Appl
41	24	82.8	26	4	US-09-299-495E-37	Sequence 37, Appl
42	24	82.8	33	2	US-08-690-011A-38	Sequence 38, Appl
43	24	82.8	33	4	US-09-299-495F-38	Sequence 38, Appl
44	24	82.8	34	2	US-08-690-011A-53	Sequence 53, Appl
45	24	82.8	34	2	US-08-690-011A-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-08-726-306A-183
Sequence 183, Application US/08726306A
Patent No. 5958684
GENERAL INFORMATION:
APPLICANT: van Leeuwen, Frederik Willem
APPLICANT: Burbach, Johannes Peter Henri
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1 Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,306A
FILING DATE: 02-Oct-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080.4
FILING DATE: 02-Oct-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/009,832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 183:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-726-306A-183
Query Match 100.0%; Score 29; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEX 6
111111
DB 112 NEALER 117

RESULT 2

US-08-792-014-1
Sequence 1, Application US/08792014
Patent No. 6063594
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: NOVEL HUMAN ANION CHANNEL
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/792.014
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0206 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
US-08-792-014-1
Query Match 89.7%; Score 26; DB 3; Length 253;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALER 6
111111
DB 137 NEALER 142

RESULT 3

US-09-443-948-1
Sequence 1, Application US/09443948
Patent No. 6228616
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: NOVEL HUMAN ANION CHANNEL
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/443,948
FILING DATE: 19-Nov. 6228616-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/792,014
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0206 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: Consensus
US-09-443-948-1
Query Match 89.7%; Score 26; DB 4; Length 253;
Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALER 6
111111
DB 137 NEALER 142

RESULT 4

US-08-961-083-218
Sequence 218, Application US/08961083
Patent No. 6159469
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,083
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:


```

      ; ORIGINAL SOURCE:
      ; ORGANISM: Mus musculus
      ; US-08-426-627-12

      Query Match
      Best Local Similarity 86.2%; Score 25; DB 1; Length 16;
      Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0

      QY 1 NEALEK 6
          ||| 11
      Db 10 NEALEK 15

      RESULT 6
      US-08-743-200-10
      ; Sequence 10, Application US/08743200
      ; Patent No. 5861260
      ; GENERAL INFORMATION:
      ; APPLICANT: Dohsey, Stephen J.
      ; TITLE OF INVENTION: DIAGNOSTIC METHODS FOR SCREENING
      ; TITLE OF INVENTION: PATIENTS FOR SCLERODERMA
      ; NUMBER OF SEQUENCES: 36
      ; CORRESPONDENCE ADDRESS:
      ; ADDRESSEE: Fish & Richardson P.C.
      ; STREET: 225 Franklin Street
      ; CITY: Boston
      ; STATE: MA
      ; COUNTRY: US
      ; ZIP: 02110-2804
      ; COMPUTER READABLE FORM:
      ; MEDIUM TYPE: Diskette
      ; COMPUTER: IBM Compatible
      ; OPERATING SYSTEM: DOS
      ; SOFTWARE: FASTSEQ Version 2.0
      ; CURRENT APPLICATION DATA:
      ; APPLICATION NUMBER: US/08/743,200
      ; FILING DATE: 05-NOV-1996
      ; PRIOR APPLICATION DATA:
      ; APPLICATION NUMBER:
      ; FILING DATE:
      ; ATTORNEY/AGENT INFORMATION:
      ; NAME: Fasse, J. Peter
      ; REGISTRATION NUMBER: 32,983
      ; REFERENCE/DOCKET NUMBER: 07917/025001
      ; TELECOMMUNICATION INFORMATION:
      ; TELEPHONE: 617-542-5070
      ; TELEFAX: 617-542-8906
      ; INFORMATION FOR SEQ ID NO: 10:
      ; SEQUENCE CHARACTERISTICS:
      ; LENGTH: 37 amino acids
      ; TYPE: amino acid
      ; STRANDEDNESS: single
      ; TOPOLOGY: linear
      ; MOLECULE TYPE: protein
      ; FRAGMENT TYPE: internal
      ; US-08-743-200-10

      Query Match
      Best Local Similarity 83.3%; Score 25; DB 2; Length 37;
      Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

      QY 1 NEALEK 6
          |||| 11
      Db 1 NEALEK 6

      RESULT 7
      US-08-426-627-15
      ; Sequence 15, Application US/08426627
      ; Patent No. 5756664
      ; GENERAL INFORMATION:
      ; APPLICANT: Atawara-Hamamoto, Yoko

```

APPLICANT: Kikuno, Reiko
APPLICANT: Takeshita, Sunao
TITLE OF INVENTION: No. 575664e1 Protein with Bone Formation
TITLE OF INVENTION: Ability and Process for Its Production.
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,627
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/036,841
FILING DATE: 25-MAR-1993
APPLICATION NUMBER: JP 4-71501
FILING DATE: 27-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hammond, Alan W.
REGISTRATION NUMBER: 35,178
REFERENCE/DOCKET NUMBER: 02481-1285-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-426-627-15

Query Match 86.2%; Score 25; DB 1; Length 135;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NEALEX 6
Db 43 NEALEX 48

RESULT 8
US-08-477-396A-2
Sequence 2, Application US/08477396A
Patent No. 5872235
GENERAL INFORMATION:
APPLICANT: Chen, Ian Bo
APPLICANT: Bao, Shideng
APPLICANT: Liu, Yuan
TITLE OF INVENTION: A NOVEL TUMOR MARKER AND NOVEL METHOD OF
TITLE OF INVENTION: ISOLATING SAME
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible.

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,396A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/146,488
FILING DATE: 29-OCT-1993
APPLICATION NUMBER: US 08/448,388
FILING DATE: 28-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/12502
FILING DATE: 31-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Helne, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-333BX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
FILING DATE: (617) 451-0313
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 212 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-396A-2

Query Match 86.2%; Score 25; DB 2; Length 212;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NEALEX 6
Db 66 NEALEX 71

RESULT 9
US-09-134-001C-3754
Sequence 3754, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3754
LENGTH: 252
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3754

Query Match 86.2%; Score 25; DB 4; Length 252;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
Db 30 NEALEX 35

RESULT 10
US-08-353-476-71
Sequence 71, Application US/08353476
Patent No. 5871902

GENERAL INFORMATION:
APPLICANT: Weininger, Susan
TITLE OF INVENTION: METHOD OF DETECTION OF DNA WITH A
TITLE OF INVENTION: SPECIFIC SEQUENCE COMPOSITION
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st St., Suite A-1
CITY: Gainesville
STATE: Florida
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,476
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Benchen, Gerard H
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: GP-100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904) 375-8100
TELEFAX: (904) 372-5800
INFORMATION FOR SEQ ID NO: 71:
SEQUENCE CHARACTERISTICS:
LENGTH: 314 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
US-08-353-476-71

Query Match 86.2%; Score 25; DB 2; Length 314;
Best Local Similarity 83.3%; Pred. No. 2.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALKR 6
DB 67 NEALKR 72

RESULT 11
US-09-134-001C-4571
Sequence 4571, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4571
LENGTH: 399
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4571

Query Match 86.2%; Score 25; DB 4; Length 399;
Best Local Similarity 83.3%; Pred. No. 3.8e+02;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NEALKR 6
DB 144 NEALKR 149

RESULT 12
US-09-362-473-12
Sequence 12, Application US/09362473
Patent No. 6218169
GENERAL INFORMATION:
APPLICANT: Cahoon, Edgar B.
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Falco, S. Carl
APPLICANT: Morgante, Michele
APPLICANT: Rafalski, J. Antoni
APPLICANT: Hiltz, William D.
APPLICANT: Kinney, Anthony J.
TITLE OF INVENTION: Aromatic Amino Acid Catabolism Enzymes
FILE REFERENCE: BB-1197
CURRENT APPLICATION NUMBER: US/09/362,473
CURRENT FILING DATE: 1999-07-28
EARLIER APPLICATION NUMBER: 60/094,783
EARLIER FILING DATE: JULY 31, 1998
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Microsoft Office 97
SEQ ID NO 12
LENGTH: 429
TYPE: PRT
ORGANISM: Oryza sativa
US-09-362-473-12

Query Match 86.2%; Score 25; DB 4; Length 429;
Best Local Similarity 83.3%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALKR 6
DB 112 NEALKR 117

RESULT 13
US-08-484-105-8
Sequence 8, Application US/08484105
Patent No. 5589341
GENERAL INFORMATION:
APPLICANT: STILLMAN, Bruce
APPLICANT: BELL, Stephen P
APPLICANT: KOBAYASHI, Ryuj1
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: MCNALLY, Francis J
APPLICANT: LAURENSEN, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,105

FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 529 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-105-8

Query Match 86.2%; Score 25; DB 1; Length 529;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NEALEK 6
DB 306 NETLEK 311

RESULT 14
US-08-484-106-8
Sequence 8, Application US/08484106
Patent No. 5614618
GENERAL INFORMATION:
APPLICANT: STILLMAN, Bruce
APPLICANT: BELL, Stephen P
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: KINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: MCNALLY, Francis J
APPLICANT: LAURENSEN, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,106
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 529 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-106-8

Query Match 86.2%; Score 25; DB 1; Length 529;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NEALEK 6
DB 306 NETLEK 311

RESULT 15
US-07-878-960-2
Sequence 2, Application US/07878960
Patent No. 5444164
GENERAL INFORMATION:
APPLICANT: Purchio, Anthony F.
APPLICANT: Skouter, John
APPLICANT: Neubauer, Michael G.
TITLE OF INVENTION: TGF-BETA INDUCED GENE AND PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bristol-Myers Squibb Company
STREET: 3005 First Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98121

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/878,960
FILING DATE: 05-MAY-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/833,835
FILING DATE: 05-FEB-1992
NAME: 530
ATTORNEY/AGENT INFORMATION:
NAME: Sorrentino, Joseph M.
REGISTRATION NUMBER: 32,598
REFERENCE/DOCKET NUMBER: ON0092-
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206/727-3601
TELEFAX: 206/728-4800

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 683 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: LUNG
CELL TYPE: ADENOCARCINOMA
CELL LINE: A549
US-07-878-960-2

Query Match 86.2%; Score 25; DB 1; Length 683;
Best Local Similarity 83.3%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NEALEK 6
DB 282 NEALEK 287

Search completed: November 13, 2002, 13:18:04
Job time : 9.68085 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comphen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:18:18 ; Search time 4.08511 Seconds
(without alignments)
22.121 Million cell updates/sec

Title: US-09-856-086-1

Sequence: 1 NEALKR 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubppaa/CTC_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubppaa/CTC_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	27	93.1	14	10	US-09-864-761-38288
2	27	93.1	94	10	US-09-815-242-5637
3	27	93.1	147	10	US-09-815-242-12398
4	27	93.1	147	10	US-09-815-242-12968
5	27	93.1	147	10	US-09-815-242-13154
6	27	93.1	886	10	US-09-801-368-180
7	26	89.7	18	10	US-09-864-761-36641
8	26	89.7	275	10	US-09-815-242-11302
9	26	89.7	423	10	US-09-815-242-5817
10	26	89.7	505	10	US-09-938-803-20
11	26	89.7	565	10	US-09-765-272-218
12	26	89.7	578	10	US-09-815-242-13023
13	25	86.2	123	10	US-09-864-761-38479
14	25	86.2	429	10	US-09-742-954-12
15	25	86.2	466	10	US-09-934-066-2
16	25	86.2	683	10	US-09-923-217-122
17	25	86.2	683	10	US-09-833-263-122
18	25	86.2	790	10	US-09-925-301-1313
19	25	86.2	995	9	US-09-486-734A-2

20	25	86.2	1127	10	US-09-815-242-5373	Sequence 5373, App
21	25	86.2	1158	10	US-09-815-242-12522	Sequence 12522, A
22	24	82.8	122	10	US-09-801-574-63	Sequence 63, Appl
23	24	82.8	148	10	US-09-795-926-10	Sequence 10, Appl
24	24	82.8	186	10	US-09-795-926-12	Sequence 12, Appl
25	24	82.8	227	10	US-09-923-302-604	Sequence 604, App
26	24	82.8	240	10	US-09-864-761-43091	Sequence 43091, A
27	24	82.8	240	10	US-09-815-242-11656	Sequence 11656, A
28	24	82.8	271	10	US-09-925-302-680	Sequence 680, App
29	24	82.8	271	10	US-09-815-242-10050	Sequence 10050, A
30	24	82.8	285	10	US-09-815-242-13933	Sequence 13933, A
31	24	82.8	293	10	US-09-968-958-2	Sequence 2, Appl
32	24	82.8	293	10	US-09-968-958-4	Sequence 4, Appl
33	24	82.8	303	10	US-09-795-926-2	Sequence 2, Appl
34	24	82.8	356	10	US-09-815-242-11523	Sequence 11523, A
35	24	82.8	362	10	US-09-815-242-11352	Sequence 11352, A
36	24	82.8	387	9	US-09-895-913A-120	Sequence 120, App
37	24	82.8	393	10	US-09-776-635-32	Sequence 32, Appl
38	24	82.8	393	10	US-09-732-384-3	Sequence 3, Appl
39	24	82.8	393	10	US-09-860-211-9	Sequence 9, Appl
40	24	82.8	394	12	US-10-155-059-4	Sequence 4, Appl
41	24	82.8	396	10	US-09-815-242-13413	Sequence 13413, A
42	24	82.8	396	10	US-09-805-847-2	Sequence 2, Appl
43	24	82.8	397	10	US-09-815-242-13591	Sequence 13591, A
44	24	82.8	401	10	US-09-924-256A-88	Sequence 88, Appl
45	24	82.8	441	10	US-09-912-020-272	Sequence 272, App

ALIGNMENTS

RESULT 1
US-09-864-761-38288
Sequence 38288, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmclca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 38288
LENGTH: 94
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL122003.7
OTHER INFORMATION: EXPRESSED IN BR474, SIGNAL - 1.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 1.2
OTHER INFORMATION: EST_HUMAN HIT: AW963611.1, EVALU 5.00e-07
OTHER INFORMATION: SWISSPROT HIT: P54156, EVALU 4.40e+00
US-09-864-761-38288

Query Match 93.1%; Score 27; DB 10; Length 94;
Best Local Similarity 83.3%; Pred. No. 12;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
Db 80 NEALEK 85

RESULT 2
US-09-815-242-5637
Sequence 5637, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5637
LENGTH: 144
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5637

Query Match 93.1%; Score 27; DB 10; Length 144;
Best Local Similarity 83.3%; Pred. No. 19;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
Db 20 NEALEK 25

RESULT 3
US-09-815-242-12398
Sequence 12398, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12398
LENGTH: 147
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12398

Query Match 93.1%; Score 27; DB 10; Length 147;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
Db 23 NEALEK 28

RESULT 4
US-09-815-242-12968
Sequence 12968, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242

```

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12968
; LENGTH: 147
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
; US-09-815-242-12968

Query Match          93.1%; Score 27; DB 10; Length 147;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
DB 23 NEALEX 28

RESULT 5
US-09-815-242-13154
; Sequence 13154, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsson, Karl L.
; APPLICANT: Zybkind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13154
; LENGTH: 147
; TYPE: PRF
; ORGANISM: Staphylococcus aureus
; US-09-815-242-13154

Query Match          93.1%; Score 27; DB 10; Length 147;
Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

Best Local Similarity 83.3%; Pred. No. 20;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
DB 23 NEALEX 28

RESULT 6
US-09-801-368-180
; Sequence 180, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 180
; LENGTH: 886
; TYPE: PRF
; ORGANISM: Saccharomyces cerevisiae
; US-09-801-368-180

Query Match          93.1%; Score 27; DB 10; Length 886;
Best Local Similarity 83.3%; Pred. No. 1,3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
DB 82 NEALEX 87

RESULT 7
US-09-864-761-36641
; Sequence 36641, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
```



```

; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 36641
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010826.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EST_HUMAN HIT: AV752442.1, EVALUO 2.20e-02
; US-09-864-761-36641

Query Match      89.7%; Score 26; DB 10; Length 18;
Best Local Similarity 83.3%; Pred. No. 3.7;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 NEALKER 6
DB      7 NEALKER 12

; RESULT 8
; US-09-815-242-11302
; Sequence 11302, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; Prokaryotes
```

```

; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11302
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Helicobacter pylori
; US-09-815-242-11302

Query Match      89.7%; Score 26; DB 10; Length 275;
Best Local Similarity 83.3%; Pred. No. 64;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 NEALKER 6
DB      41 NEALKER 46

; RESULT 9
; US-09-815-242-5817
; Sequence 5817, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5817
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE: \
```

NAME/KEY: VARIANT
LOCATION: (1)...(423)
OTHER INFORMATION: Xaa - Any Amino Acid
US-09-815-242-5817

Query Match
Best Local Similarity 89.7%; Score 26; DB 10; Length 423;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
1:|||||
DB 95 NDALX 100

RESULT 10
US-09-938-803-20
Sequence 20, Application US/09938803
Patent No. US20020076762A1
GENERAL INFORMATION:
APPLICANT: Yue, Henry
APPLICANT: Tang, Y. Tom
APPLICANT: Lal, Preethi
APPLICANT: Reddy, Rocco
APPLICANT: Baughn, Mariah R.
APPLICANT: Yang, Junming
APPLICANT: Azimzal, Yalda
TITLE OF INVENTION: FULL-LENGTH EXPRESSED GENETIC MARKERS
FILE REFERENCE: PF-0695 US
CURRENT APPLICATION NUMBER: US/09/938, 803
CURRENT FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 09/311,894
PRIOR FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PERL Program
SEQ ID NO 20
LENGTH: 505
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte Clone 3039890
US-09-938-803-20

Query Match
Best Local Similarity 89.7%; Score 26; DB 10; Length 505;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
1:|||||
DB 39 NDALX 44

RESULT 11
US-09-765-272-218
Sequence 218, Application US/09765272
Patent No. US20020061545A1
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 452
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44M storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 218:
SEQUENCE CHARACTERISTICS:
LENGTH: 565 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 218:
US-09-765-272-218

Query Match
Best Local Similarity 89.7%; Score 26; DB 10; Length 565;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
1:|||||
DB 357 NEALEX 362

RESULT 12
US-09-815-242-13023
Sequence 13023, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA, 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13023
LENGTH: 578
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-13023

Query Match
89.7%; Score 26; DB 10; Length 578;

Best Local Similarity 83.3%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALER 6
|:||||
Db 242 NDALER 247

RESULT 13

US-09-864-761-38479
; Sequence 38479, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmlca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38479
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006385.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.2

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6
US-09-864-761-38479

Query Match 86.2%; Score 25; DB 10; Length 123;
Best Local Similarity 83.3%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NEALER 6
|:||||
Db 50 NETLER 55

RESULT 14

US-09-742-954-12
; Sequence 12, Application US/09742954
; Patent No. US20010005749A1
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Morgante, Michele
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; TITLE OF INVENTION: Aromatic Amino Acid Catabolism Enzymes
; FILE REFERENCE: BB-1197
; CURRENT APPLICATION NUMBER: US/09/742,954
; CURRENT FILING DATE: 2000-12-21
; PRIOR FILING DATE: 60/094,783
; PRIOR FILING DATE: JULY 31, 1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-09-742-954-12

Query Match 86.2%; Score 25; DB 10; Length 429;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALER 6
|:||||
Db 112 NEALKK 117

RESULT 15

US-09-934-066-2
; Sequence 2, Application US/09934066
; Patent No. US20020108149A1
; GENERAL INFORMATION:
; APPLICANT: Jung, Darren B.
; APPLICANT: Jung, Rudolf
; TITLE OF INVENTION: Methods of Increasing Polypeptide
; FILE REFERENCE: 35718/237251
; CURRENT APPLICATION NUMBER: US/09/934,066
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,804
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-09-934-066-2

Query Match 86.2%; Score 25; DB 10; Length 466;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NEALEK 6
| | | |
Db 184 NEVLEK 189

Search completed: November 13, 2002, 13:40:17
Job time : 4.22796 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:16:18 ; Search time 120.383 seconds
(without alignments)
32.134 Million cell updates/sec

Title: US-09-856-086-1

Perfect score: 29

Sequence: 1 NEALEK 6

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending-Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US110_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US111_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	35	US-09-914-259-5	Sequence 5, Appl1
2	29	100.0	92	US-09-252-691-9015	Sequence 9015, Ap
3	29	100.0	92	US-09-252-691C-9015	Sequence 9015, Ap
4	29	100.0	93	US-10-234-432-84	Sequence 84, Appl
5	29	100.0	128	US-09-791-537-130121	Sequence 130121,
6	29	100.0	130	US-09-791-537-130123	Sequence 130123,

7	29	100.0	132	US-09-791-537-129991	Sequence 129991,
8	29	100.0	132	US-09-791-537-129992	Sequence 129992,
9	29	100.0	132	US-09-791-537-129993	Sequence 129993,
10	29	100.0	132	US-09-791-537-130012	Sequence 130012,
11	29	100.0	132	US-09-791-537-130014	Sequence 130014,
12	29	100.0	133	US-09-791-537-135781	Sequence 135781,
13	29	100.0	133	US-09-791-537-135782	Sequence 135782,
14	29	100.0	133	US-09-791-537-135783	Sequence 135783,
15	29	100.0	133	US-09-791-537-135784	Sequence 135784,
16	29	100.0	133	US-09-791-537-135785	Sequence 135785,
17	29	100.0	133	US-09-791-537-135787	Sequence 135787,
18	29	100.0	133	US-09-791-537-135788	Sequence 135788,
19	29	100.0	133	US-09-791-537-135789	Sequence 135789,
20	29	100.0	133	US-09-791-537-135790	Sequence 135790,
21	29	100.0	133	US-09-791-537-135791	Sequence 135791,
22	29	100.0	133	US-09-791-537-135792	Sequence 135792,
23	29	100.0	133	US-09-791-537-135793	Sequence 135793,
24	29	100.0	133	US-09-791-537-135794	Sequence 135794,
25	29	100.0	133	US-09-791-537-135795	Sequence 135795,
26	29	100.0	133	US-09-791-537-135796	Sequence 135796,
27	29	100.0	133	US-09-791-537-135797	Sequence 135797,
28	29	100.0	133	US-09-791-537-135798	Sequence 135798,
29	29	100.0	133	US-09-791-537-135799	Sequence 135799,
30	29	100.0	133	US-09-791-537-135840	Sequence 135840,
31	29	100.0	133	US-09-791-537-150723	Sequence 150723,
32	29	100.0	141	US-09-834-366-15179	Sequence 15179, A
33	29	100.0	141	US-60-197-873-15179	Sequence 15179, A
34	29	100.0	149	US-09-107-532A-5296	Sequence 5296, Ap
35	29	100.0	149	US-09-107-532A-5296	Sequence 5296, Ap
36	29	100.0	149	US-60-202-188-120	Sequence 120, App
37	29	100.0	142	US-09-791-537-64354	Sequence 64354, A
38	29	100.0	142	US-09-791-537-76695	Sequence 76695, A
39	29	100.0	180	US-09-791-537-63072	Sequence 63072, A
40	29	100.0	180	US-09-791-537-108547	Sequence 108547,
41	29	100.0	182	US-60-162-357-962	Sequence 962, App
42	29	100.0	183	US-09-292-085-5	Sequence 5, Appl1
43	29	100.0	183	US-09-791-537-118316	Sequence 118316,
44	29	100.0	185	US-09-791-537-61085	Sequence 61085, A
45	29	100.0	220	US-09-513-996A-55574	Sequence 55574, A

ALIGNMENTS

RESULT 1
US-09-914-259-5
Sequence 5, Application US/09914259
APPLICANT: Makowski, Lee
APPLICANT: Hyman, Paul
TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
FILE REFERENCE: 8471-010-999
CURRENT FILING DATE: 2000-11-21
NUMBER OF SEQ ID NOS: 180
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 35
TYPE: PRT
ORGANISM: Homo sapiens
US-09-914-259-5

Query Match 100.0%; Score 29; DB 23; Length 35;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEK 6
DB 8 NEALEK 13

RESULT 2

```
US-09-252-691-9015
; Sequence 9015, Application US/09252691B
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252.691B
; CURRENT FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 11324
; SEQ ID NO 9015
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-09-252-691-9015

Query Match          100.0%; Score 29; DB 16; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
Db 15 NEALEX 20

RESULT 3
US-09-252-691C-9015
; Sequence 9015, Application US/09252691C
; GENERAL INFORMATION:
; APPLICANT: Keith G. Weinstein et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
; TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.135
; CURRENT APPLICATION NUMBER: US/09/252.691C
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/094,145
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: US 60/074,787
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 11326
; SEQ ID NO 9015
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-09-252-691C-9015

Query Match          100.0%; Score 29; DB 16; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
Db 15 NEALEX 20

RESULT 4
US-10-234-432-84
; Sequence 84, Application US/10234432
; GENERAL INFORMATION:
; APPLICANT: Homer, Mary J.
; APPLICANT: Lodges, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Persing, David H.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF BABESIA INFECTION
; FILE REFERENCE: 210121.560
; CURRENT APPLICATION NUMBER: US/10/234.432
; CURRENT FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 84
; LENGTH: 93

; TYPE: PRT
; ORGANISM: Babesia sp. WAI
US-10-234-432-84

Query Match          100.0%; Score 29; DB 26; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
Db 30 NEALEX 35

RESULT 5
US-09-791-537-130121
; Sequence 130121, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 130121
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-130121

Query Match          100.0%; Score 29; DB 21; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
Db 70 NEALEX 75

RESULT 6
US-09-791-537-130123
; Sequence 130123, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 130123
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-130123

Query Match          100.0%; Score 29; DB 21; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
Db 72 NEALEX 77

RESULT 7
US-09-791-537-129991
```

```
; Sequence 129991, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 129991
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-129991

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 132;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEX 6
Db 74 NEALEX 79

; Sequence 129992, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 129992
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-129992

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 132;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEX 6
Db 74 NEALEX 79

; Sequence 129993, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 129993
; LENGTH: 132
; TYPE: PRT
US-09-791-537-129993

; ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-129993

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 132;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEX 6
Db 74 NEALEX 79

; Sequence 130012, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 130012
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-130012

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 132;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEX 6
Db 74 NEALEX 79

; Sequence 130014, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 130014
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-130014

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 132;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEX 6
Db 75 NEALEX 80

; Sequence 135781, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 135781
; LENGTH: 132
; TYPE: PRT
US-09-791-537-135781

; ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-135781

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 132;
Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEX 6
Db 75 NEALEX 80
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Bionomix, Inc.
/ APPLICANT: Debe, Derek
/ APPLICANT: Danzer, Joseph
/ TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
/ FILE REFERENCE: 261/210
/ CURRENT APPLICATION NUMBER: US/09/791,537
/ CURRENT FILING DATE: 2001-02-22
/ NUMBER OF SEQ ID NOS: 153055
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 135781
/ LENGTH: 133
/ TYPE: PRT
/ ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-135781

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 133;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEK 6
Db 75 NEALEK 80

RESULT 13
US-09-791-537-135782
/ Sequence 135782, Application US/09791537
/ GENERAL INFORMATION:
/ APPLICANT: Bionomix, Inc.
/ APPLICANT: Debe, Derek
/ APPLICANT: Danzer, Joseph
/ TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
/ FILE REFERENCE: 261/210
/ CURRENT APPLICATION NUMBER: US/09/791,537
/ CURRENT FILING DATE: 2001-02-22
/ NUMBER OF SEQ ID NOS: 153055
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 135782
/ LENGTH: 133
/ TYPE: PRT
/ ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-135782

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 133;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEK 6
Db 75 NEALEK 80

RESULT 14
US-09-791-537-135783
/ Sequence 135783, Application US/09791537
/ GENERAL INFORMATION:
/ APPLICANT: Bionomix, Inc.
/ APPLICANT: Debe, Derek
/ APPLICANT: Danzer, Joseph
/ TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
/ FILE REFERENCE: 261/210
/ CURRENT APPLICATION NUMBER: US/09/791,537
/ CURRENT FILING DATE: 2001-02-22
/ NUMBER OF SEQ ID NOS: 153055
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 135783
/ LENGTH: 133
/ TYPE: PRT
/ ORGANISM: Human Immunodeficiency virus type 1
```

```
US-09-791-537-135783

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 133;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEK 6
Db 75 NEALEK 80

RESULT 15
US-09-791-537-135784
/ Sequence 135784, Application US/09791537
/ GENERAL INFORMATION:
/ APPLICANT: Bionomix, Inc.
/ APPLICANT: Debe, Derek
/ APPLICANT: Danzer, Joseph
/ TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY M
/ FILE REFERENCE: 261/210
/ CURRENT APPLICATION NUMBER: US/09/791,537
/ CURRENT FILING DATE: 2001-02-22
/ NUMBER OF SEQ ID NOS: 153055
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 135784
/ LENGTH: 133
/ TYPE: PRT
/ ORGANISM: Human Immunodeficiency virus type 1
US-09-791-537-135784

Query Match
Best Local Similarity 100.0%; Score 29; DB 21; Length 133;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEK 6
Db 75 NEALEK 80
```

Search completed: November 13, 2002, 13:39:13
Job time : 121.383 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2002, 13:17:03 : Search time 2.42553 Seconds
(without alignments)
29,431 Million cell updates/sec

Title: US-09-856-086-1

Perfect score: 29

Sequence: 1 NEALEX 6

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 41632 seqs, 11897606 residues

Total number of hits satisfying chosen parameters: 41632

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database: Pending_Patents_AA_New:*

1: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*

2: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*

3: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*

4: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*

5: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*

6: /cgn2_6/ptodata/1/paa/US11_NEW_COMB.pep:*

7: /cgn2_6/ptodata/1/paa/US12_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	6	5	US-09-856-086-1
2	29	100.0	103	5	US-09-513-999C-6211
3	26	89.7	118	5	US-09-513-999C-5017
4	25	86.2	105	5	US-09-513-999C-6681
5	25	86.2	252	6	US-10-092-411A-3754
6	25	86.2	399	6	US-10-092-411A-4571
7	25	86.2	636	6	US-10-204-752A-33
8	24	82.8	119	5	US-09-513-999C-5465
9	24	82.8	181	6	US-10-092-411A-2892
10	24	82.8	181	6	US-10-092-411A-5665
11	24	82.8	235	6	US-10-092-411A-5142
12	24	82.8	244	6	US-10-092-411A-4784
13	24	82.8	393	6	PCT-US02-31461-3
14	24	82.8	393	6	US-10-274-874-4
15	24	82.8	410	6	US-10-092-411A-3927
16	24	82.8	458	6	US-10-092-411A-4663
17	24	82.8	493	6	US-10-092-411A-4388
18	24	82.8	652	6	PCT-US02-32851-17
19	24	82.8	876	6	US-10-092-411A-3943
20	23	79.3	29	5	US-09-856-086-3
21	23	79.3	29	6	US-10-188-444-9
22	23	79.3	29	6	US-10-188-444-9
23	23	79.3	29	6	US-10-103-597A-8
24	23	79.3	29	6	US-10-103-597A-9
25	23	79.3	90	1	PCT-US02-21946A-4
26	23	79.3	137	5	US-09-513-999C-6185

ALIGNMENTS

27	23	79.3	150	6	US-10-113-709A-4	Sequence 4, Appl
28	23	79.3	197	6	US-10-264-237-2423	Sequence 2423, Ap
29	23	79.3	274	6	US-10-092-411A-2995	Sequence 2995, Ap
30	23	79.3	307	6	US-10-092-411A-5144	Sequence 5144, Ap
31	23	79.3	323	6	US-10-274-694-17	Sequence 17, Appl
32	23	79.3	340	6	US-10-264-213-141	Sequence 141, Appl
33	23	79.3	340	6	US-10-264-213-252	Sequence 252, App
34	23	79.3	347	6	US-10-264-237-2401	Sequence 2401, Ap
35	23	79.3	454	6	US-10-092-411A-4438	Sequence 4438, Ap
36	23	79.3	531	6	US-10-092-411A-3574	Sequence 3574, Ap
37	23	79.3	621	6	US-10-131-813A-40	Sequence 40, Appl
38	23	79.3	621	6	US-10-131-813A-40	Sequence 40, Appl
39	23	79.3	621	6	US-10-131-823A-40	Sequence 40, Appl
40	23	79.3	621	6	US-10-131-824A-40	Sequence 40, Appl
41	23	79.3	621	6	US-10-131-826A-40	Sequence 40, Appl
42	23	79.3	621	6	US-10-131-829A-40	Sequence 40, Appl
43	23	79.3	621	6	US-10-125-926A-40	Sequence 40, Appl
44	23	79.3	621	6	US-10-127-829A-40	Sequence 40, Appl
45	23	79.3	621	6	US-10-127-831A-40	Sequence 40, Appl

RESULT 1
US-09-856-086-1
Sequence 1, Application US/09856086
GENERAL INFORMATION:
APPLICANT: EBRINGER, ALAN
TITLE OF INVENTION: DIAGNOSIS OF DEMENTIATING OR SPONGIFORM DISEASE
FILE REFERENCE: 78104.040
CURRENT APPLICATION NUMBER: US/09/856, 086
CURRENT FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 6
TYPE: PRT
ORGANISM: Homo sapiens, Bos
US-09-856-086-1

Query Match
Best local similarity 100.0%; Pred. No. 3.8e+04;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
Db 1 NEALEX 6

RESULT 2
US-09-513-999C-6211
Sequence 6211, Application US/09513999C
GENERAL INFORMATION:
APPLICANT: Dumas Mline Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: 59, US2, REG
CURRENT APPLICATION NUMBER: US/09/513, 999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 6211
LENGTH: 103
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 90
OTHER INFORMATION: Xaa-Leu or Val

US-09-513-999C-6211

Query Match 100.0%; Score 29; DB 5; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.3; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

OY 1 NEALEK 6
DB 80 NEALEK 85

RESULT 3

US-09-513-999C-5017
Sequence 5017, Application US/09513999C
GENERAL INFORMATION:
APPLICANT: Dumas Mline Edwards, J.B.
APPLICANT: Duclet, A.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: 59, US2, REG
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent, pm
SEQ ID NO 5017
LENGTH: 118
TYPE: PRT
ORGANISM: Homo sapiens
US-09-513-999C-5017

Query Match 89.7%; Score 26; DB 5; Length 118;
Best Local Similarity 83.3%; Pred. No. 12; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0;

OY 1 NEALEK 6
DB 39 NDALK 44

RESULT 4

US-09-513-999C-6681
Sequence 6681, Application US/09513999C
GENERAL INFORMATION:
APPLICANT: Dumas Mline Edwards, J.B.
APPLICANT: Duclet, A.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: 59, US2, REG
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent, pm
SEQ ID NO 6681
LENGTH: 105
TYPE: PRT
ORGANISM: Homo sapiens
US-09-513-999C-6681

Query Match 86.2%; Score 25; DB 5; Length 105;
Best Local Similarity 83.3%; Pred. No. 19; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0;

OY 1 NEALEK 6
DB 55 NEALK 60

RESULT 5
US-10-092-411A-3754

Sequence 3754, Application US/10092411A

GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-101
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5676
SEQ ID NO 3754
LENGTH: 252
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-092-411A-3754

Query Match 86.2%; Score 25; DB 6; Length 252;
Best Local Similarity 83.3%; Pred. No. 46; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0;

OY 1 NEALEK 6
DB 30 NEALEK 35

RESULT 6

US-10-092-411A-4571
Sequence 4571, Application US/10092411A
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 032796-101
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5676
SEQ ID NO 4571
LENGTH: 399
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-092-411A-4571

Query Match 86.2%; Score 25; DB 6; Length 399;
Best Local Similarity 83.3%; Pred. No. 73; Indels 0; Gaps 0;
Matches 5; Conservative 1; Mismatches 0;

OY 1 NEALEK 6
DB 144 NEYLER 149

RESULT 7

US-10-204-752A-33
Sequence 33, Application US/10204752A
GENERAL INFORMATION:
APPLICANT: OHNANI, Noriko
APPLICANT: MATSUI, Keiko
APPLICANT: YOSHIDA, Nei
APPLICANT: SUGITA, Yuji
APPLICANT: IZUHARA, Kenji
TITLE OF INVENTION: METHOD OF TESTING FOR ALLERGIC DISEASE
FILE REFERENCE: SHZ-009US

```
; CURRENT APPLICATION NUMBER: US/10/204,752A
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: JP 2000-396166
; PRIOR FILING DATE: 2000-12-26
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 33
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-204-752A-33
```

```
Query Match
Best Local Similarity 86.2%; Score 25; DB 6; Length 836;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 NEALEX 6
DB 276 NEALEX 281
```

```
RESULT 8
US-09-513-999C-5465
; Sequence 5465, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Mline Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5465
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: UNSURE
; LOCATION: 106
; OTHER INFORMATION: Xaa- * or Phe or Leu or Tyr
US-09-513-999C-5465
```

```
Query Match
Best Local Similarity 82.8%; Score 24; DB 5; Length 119;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 NEALEX 5
DB 66 NEALEX 70
```

```
RESULT 9
US-10-092-411A-2892
; Sequence 2892, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 2892
```

```
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-2892
```

```
Query Match
Best Local Similarity 82.8%; Score 24; DB 6; Length 181;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 NEALEX 6
DB 81 NEALEX 86
```

```
RESULT 10
US-10-092-411A-5665
; Sequence 5665, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 5665
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-5665
```

```
Query Match
Best Local Similarity 82.8%; Score 24; DB 6; Length 181;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 NEALEX 6
DB 123 NEALEX 128
```

```
RESULT 11
US-10-092-411A-5142
; Sequence 5142, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 5142
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-5142
```

```
Query Match
Best Local Similarity 82.8%; Score 24; DB 6; Length 235;
Best Local Similarity 83.3%; Pred. No. 71;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

OY 1 NEALEK 6
 Db 139 NEALEK 144

RESULT 12

US-10-092-411A-4784
 ; Sequence 4784, Application US/10092411A
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; FILE REFERENCE: 032796-101
 ; CURRENT APPLICATION NUMBER: US/10/092,411A
 ; CURRENT FILING DATE: 2002-03-07
 ; PRIOR APPLICATION NUMBER: US 09/134,001
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5676
 ; SEQ ID NO 4784
 ; LENGTH: 244
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-10-092-411A-4784

Query Match

Best Local Similarity 82.8%; Score 24; DB 6; Length 244;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 Db 131 NEALEK 136

RESULT 13

PCT-US02-21461-3
 ; Sequence 3, Application PC/TUS0221461
 ; GENERAL INFORMATION:
 ; APPLICANT: Eli Lilly Pharmaceuticals, Inc. et al.
 ; TITLE OF INVENTION: STR2 ACTIVITY
 ; FILE REFERENCE: 13407-015W01
 ; CURRENT APPLICATION NUMBER: PCT/US02/21461
 ; CURRENT FILING DATE: 2002-07-08
 ; PRIOR APPLICATION NUMBER: US 10/190,159
 ; PRIOR FILING DATE: 2002-07-05
 ; PRIOR APPLICATION NUMBER: US 60/303,370
 ; PRIOR FILING DATE: 2001-07-06
 ; PRIOR APPLICATION NUMBER: US 60/303,456
 ; PRIOR FILING DATE: 2001-07-06
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 393
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US02-21461-3

Query Match

Best Local Similarity 82.8%; Score 24; DB 1; Length 393;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALE 5
 Db 345 NEALE 349

RESULT 14
 US-10-274-874-4
 ; Sequence 4, Application US/10274874
 ; GENERAL INFORMATION:

APPLICANT: Trink, Barry
 APPLICANT: Jen, Jin
 APPLICANT: Ratovitski, Edward
 APPLICANT: Sidransky, David
 TITLE OF INVENTION: p40 Protein Acts as an Oncogene
 FILE REFERENCE: 01107,79765
 CURRENT APPLICATION NUMBER: US/10/274,874
 CURRENT FILING DATE: 2002-10-22
 PRIOR APPLICATION NUMBER: US/09/277,196
 PRIOR FILING DATE: 1999-03-26
 PRIOR APPLICATION NUMBER: 60/079736
 PRIOR FILING DATE: 1998-03-27
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 4
 LENGTH: 393
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-274-874-4

Query Match

Best Local Similarity 82.8%; Score 24; DB 6; Length 393;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALE 5
 Db 345 NEALE 349

RESULT 15

US-10-092-411A-3927
 ; Sequence 3927, Application US/10092411A
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCO
 ; FILE REFERENCE: 032796-101
 ; CURRENT APPLICATION NUMBER: US/10/092,411A
 ; CURRENT FILING DATE: 2002-03-07
 ; PRIOR APPLICATION NUMBER: US 09/134,001
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5676
 ; SEQ ID NO 3927
 ; LENGTH: 410
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-10-092-411A-3927

Query Match

Best Local Similarity 82.8%; Score 24; DB 6; Length 410;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 Db 200 NEALEK 205

Search completed: November 13, 2002, 13:39:38
 Job time : 3.42553 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2002, 13:13:25 ; Search time 9.31915 seconds
(without alignments)
61.895 Million cell updates/sec

Title: US-09-856-086-1

Perfect score: 29
Sequence: 1 NEALEK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	162	2 S00132	beta-lactoglobulin
2	29	100.0	180	1 LGSF	beta-lactoglobulin
3	29	100.0	183	2 A10296	translational Initia
4	29	100.0	185	2 T06912	H+-transporting tw
5	29	100.0	207	2 B81256	phosphoribosyl-AMP
6	29	100.0	247	2 F75486	uracil-DNA N-glyco
7	29	100.0	251	2 H69441	carboxylesterase (
8	29	100.0	260	2 E70467	hypothetical prote
9	29	100.0	265	2 G71032	hypothetical prote
10	29	100.0	287	2 A21762	neurofilament trip
11	29	100.0	389	2 AD1614	Acetyl-CoA acetyl
12	29	100.0	413	2 A86814	phosphoribosylamin
13	29	100.0	492	2 AH1030	probable exported
14	29	100.0	543	1 QFMSL	neurofilament trip
15	29	100.0	544	2 S07144	neurofilament trip
16	29	100.0	548	1 QFPGI	neurofilament trip
17	29	100.0	553	2 B88072	protein ZK1240.2 f
18	29	100.0	554	2 JN0094	neurofilament prot
19	29	100.0	556	2 B46024	neurofilament-L su
20	29	100.0	652	2 G82401	methyl-accepting c
21	29	100.0	868	2 G71691	hypothetical prote
22	29	100.0	1198	2 T34340	dynein heavy chain
23	29	93.1	99	2 A31048	phosphotransferase
24	27	93.1	99	2 B90518	conserved hypother
25	27	93.1	130	2 G86540	CT255 hypothetical
26	27	93.1	130	2 H72081	mazg protein, prob
27	27	93.1	144	1 H70088	phosphotransferase
28	27	93.1	144	2 B90011	PTS system, manit
29	27	93.1	251	2 D69861	RNA polymerase slg

30	27	93.1	254	2 A10309	gluconate 5-dehydr
31	27	93.1	366	2 F89936	conserved hypother
32	27	93.1	372	2 S24996	phosphopyruvate hy
33	27	93.1	407	2 T48280	hypothetical prote
34	27	93.1	449	2 D69476	TBP-interacting pr
35	27	93.1	708	2 B86470	protein -
36	27	93.1	787	2 H90543	conserved hypother
37	27	93.1	886	1 RGYL3	regulatory protein
38	26	89.7	57	2 F81308	probable coiled-co
39	26	89.7	86	2 F90490	hypothetical prote
40	26	89.7	109	2 G70317	conserved hypother
41	26	89.7	137	2 S52360	superoxide dismuta
42	26	89.7	137	2 S52362	superoxide dismuta
43	26	89.7	141	1 HAGSDC	hemoglobin alpha-D
44	26	89.7	142	2 S70484	R543 protein - rat
45	26	89.7	180	2 AE3163	hypothetical prote

ALIGNMENTS

RESULT 1

S00132

beta-lactoglobulin B - mouflon

C/Species: Ovis orientalis musimon, Ovis ammon musimon (mouflon)

C/Date: 30-Jun-1989 #sequence_rev15ion 30-Jun-1989 #text_change 12-May-1995

C/Accession: S00132

R:Godoyac-Zimmermann, J.; Confi, A.; Napolitano, L.

Biol. Chem. Hoppe-Seyler 368, 1133-1139, 1987

A/Title: The complete amino-acid sequence of dimeric beta-lactoglobulin from mouflon

A/Reference number: S00132; MWID:88106996; PMID:3426802

A/Accession: S00132

A/Molecule type: protein

A/Residues: 1-162 <GOD>

C/Superfamily: lipocalin; lipocalin homology

C/Keywords: dimer; milk

F:10-160/Domain: lipocalin homology <LIP>

F:66-160/106-119/Disulfide bonds: #status predicted

Query Match 100.0%; Score 29; DB 2; Length 162;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6

DB 130 NEALEK 135

RESULT 2

LGSF

beta-lactoglobulin precursor - sheep

N/Alternate names: beta-lactoglobulin A; beta-lactoglobulin B; beta-lactoglobulin C;

C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C/Date: 11-May-1985 #sequence_rev15ion 19-Apr-1996 #text_change 22-Jun-1999

C/Accession: J00749; J00748; A30011; S02136; A25136; A03221; S04955

R:Ali, S.; McLennaghan, M.; Simons, J.P.; Clark, A.J.

Gene 91, 201-207, 1990

A/Title: Characterisation of the alleles encoding ovine beta-lactoglobulins A and B.

A/Reference number: J00748; MWID:91007276; PMID:1976573

A/Accession: J00749

A/Molecule type: DNA

A/Residues: 1-180 <ALIB>

A/Cross-references: GB:M32232

A/Experimental source: beta-lactoglobulin B

A/Accession: J00748

A/Molecule type: DNA

A/Residues: 1-37, 'Y', '39-180 <ALIA>

A/Cross-references: GB:M32232

A/Experimental source: beta-lactoglobulin A

R:Ali, S.; Clark, A.J.

J. Mol. Biol. 199, 415-426, 1988

A/Title: Characterization of the gene encoding ovine beta-lactoglobulin. Similarity t

A:Accession: A30011
A:Molecule type: DNA
A:Residues: 1-180 <AL11>
A:Cross-references: GB:X14971
A:Experimental source: beta-lactoglobulin I
A:Accession: B30011
A:Molecule type: DNA
A:Residues: 1-37, 'Y', 39-102, 'N', 104-180 <AL12>
A:Cross-references: GB:X07009
A:Experimental source: beta-lactoglobulin II
R:Harriis, S.; Ali, S.; Anderson, S.; Archibald, A.L.; Clark, A.J.
Nucleic Acids Res. 16, 10379-10380, 1988
A:Title: Complete nucleotide sequence of the genomic ovine beta-lactoglobulin gene.
A:Reference number: S02136; MUID:89057492; PMID:3194215
A:Accession: S02136
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-180 <HAR>
A:Cross-references: EMBL:X12817; NID:g1313; PIDN:CAA31305.1; PID:g1314
R:Gaye, P.; Hue-Delahaie, D.; Mercier, J.C.; Soulier, S.; Villette, J.L.; Furet, J.P.
Biochimie 68, 1097-1107, 1986
A:Title: Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels dur
A:Reference number: A25136; MUID:87049827; PMID:3096387
A:Accession: A25136
A:Molecule type: mRNA
A:Residues: 1-180 <GAY>
A:Cross-references: GB:X04520; NID:g1315; PIDN:CAA28204.1; PID:g1316
R:Preaux, G.; Brantlitz, G.; Kolde, H.J.
Arch. Int. Physiol. Biochim. 88, B45-B46, 1980
A:Title: Primary structure of ovine beta-lactoglobulin.
A:Reference number: A03221; MUID:80219294; PMID:6155855
A:Accession: A03221
A:Molecule type: Protein
A:Residues: 19-37, 'Y', 39-180 <PRE>
R:Erhardt, G.; Godovac-Zimmermann, J.; Contl, A.
Biochem. Chem. Hoppe-Seyler 370, 757-762, 1989
A:Title: Isolation and complete primary sequence of a new ovine wild-type beta-lactoglob
A:Reference number: S04955; MUID:89374823; PMID:2775495
A:Accession: S04955
A:Molecule type: Protein
A:Residues: 19-37, 'Y', 39-165, 'O', 167-180 <ERH>
A:Experimental source: beta-lactoglobulin C
A:Comment: This protein is the major milk whey protein of ruminants and is produced in t
C:Comment: Under physiological conditions beta-lactoglobulin exists as an equilibrium m
C:Genetics:
A:Gene: BLG
A:Intons: 32/3; 79/2; 104/1; 141/1; 176/1
C:Superfamily: 11pocalin; 11pocalin homology
C:Keywords: milk; polymorphism
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-160/Product: beta-lactoglobulin #status experimental <MAT>
F:28-178/Domain: 11pocalin homology <LIP>
F:84-178,124-137/Disulfide bonds: #status predicted

Query Match 100.0%; Score 29; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
DB 148 NEALEK 153

RESULT 3
A:Accession: A10296
A:Molecule type: DNA
A:Residues: 1-207 <PAR>
A:Cross-references: GB:AL139079; GB:AL111168; NID:g6568971; PIDN:CA873592.1; PID:g6
C:Date: 02-Nov-2001 #sequence, revision 02-Nov-2001 #text, change 09-Nov-2001
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tilball, R.W.; Holden, M.T.G.; Prentice, M.B.
demo-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: A50001; MUID:21470413; PMID:11586360
A:Accession: A10296
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-183 <KUR>
A:Cross-references: GB:AL590842; PIDN:CA91237.1; PID:g15980426; GSPDB:GN00175
C:Genetics:
A:Gene: infC
C:Superfamily: translation initiation factor IF-3

Query Match 100.0%; Score 29; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
DB 40 NEALEK 45

RESULT 4
A:Accession: T06912
A:Molecule type: DNA
A:Residues: 1-185 <STR>
A:Cross-references: EMBL:U030821; NID:g1016083; PIDN:AA81255.1; PID:g1016168
C:Genetics:
A:Gene: atpF
A:Genome: cyanelle
C:Superfamily: H+-transporting ATP synthase chain I
C:Keywords: ATP biosynthesis; cyanelle; hydrolase; membrane-associated complex; thy

Query Match 100.0%; Score 29; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
DB 77 NEALEK 82

RESULT 5
A:Accession: B81256
A:Molecule type: DNA
A:Residues: 1-207 <PAR>
A:Cross-references: GB:AL139079; GB:AL111168; NID:g6568971; PIDN:CA873592.1; PID:g6
C:Date: 31-Mar-2000 #sequence, revision 31-Mar-2000 #text, change 16-Aug-2002
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelsey, J.M.; Churcher, C.; Basham, D.; Ch
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; B
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveal
A:Reference number: A81250; MUID:20150912; PMID:10688204
A:Accession: B81256
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-207 <PAR>
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: hlsI; Cj1604
C:Superfamily: hlsI bifunctional enzyme; hlsI bifunctional enzyme homology; hlsI pr
C:Keywords: hydrolase

Query Match 100.0%; Score 29; DB 2; Length 207;

Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 40 NEALEK 45

RESULT 6

urc11-DNA N-glycosylase - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C/Accession: F75486

R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A>Title: genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A/Reference number: A75250; MID:20036896; PMID:10567266

A/Accession: F75486

A>Status: preliminary

A/Molecule type: DNA

A/Residues: 1247 <NHI>

A/Cross-references: GB:AE001926; GB:AE000513; NID:96458395; PIDN:AAF10269.1; PID:9645839

A/Experimental source: strain R1

C/Genetics:

A/Map position: 1

C/Superfamily: urc11-DNA glycosylase

Query Match 100.0%; Score 29; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 51;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 225 NEALEK 230

RESULT 7

H69441 carboxylesterase (est-2) homolog - *Archaeoglobus fulgidus*

C/Species: *Archaeoglobus fulgidus*

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Feb-2000

C/Accession: H69441

R/Blank, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

., Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Nature 390, 364-370, 1997

A/Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kahne, B.P.; Sykes, S.

Smith, H.O.; Moose, C.R.; Venter, J.C.

A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A/Reference number: A69250; MID:98049343; PMID:9389475

A/Accession: H69441

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-251 <RLF>

A/Cross-references: GB:AE000996; GB:AE000782; NID:92689319; PIDN:AA89709.1; PID:9264902

C/Superfamily: peroxidase

Query Match 100.0%; Score 29; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 52;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 239 NEALEK 244

RESULT 8

E70467 hypothetical protein ag_1956 - *Aquifex aeolicus*

C/Species: *Aquifex aeolicus*

C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 08-Sep-2000

C/Accession: E70467

R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;

V.

Nature 392, 353-358, 1998

A>Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.

A/Reference number: A70300; MID:98196666; PMID:9537320

A/Accession: E70467

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-260 <AOF>

A/Cross-references: GB:AE000765; NID:92984199; PIDN:AA07740.1; PID:92984210; GB:AE00

A/Experimental source: strain VFS

C/Genetics:

A/Map position: 1956

C/Superfamily: *Aquifex aeolicus* hypothetical protein ag_1956

Query Match 100.0%; Score 29; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 54;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 45 NEALEK 50

RESULT 9

G71032 hypothetical protein PH151 - *Pyrococcus horikoshii*

C/Species: *Pyrococcus horikoshii*

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000

C/Accession: G71032

R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hatake, Y.; Hino, Y.; Yamamoto, S.; Se

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kusuda, N.; Ogu

DNA Res. 5, 55-76, 1998

A>Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A/Reference number: A71000; MID:98344137; PMID:9679194

A/Accession: G71032

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-265 <KAM>

A/Cross-references: GB:AP000006; NID:93236133; PIDN:BA030663.1; PID:93257980

A/Experimental source: strain OT3

A/Note: this accession replaces an interim accession for a sequence replaced by Genba

C/Genetics:

A/Map position: PH151

Query Match 100.0%; Score 29; DB 2; Length 265;
Best Local Similarity 100.0%; Pred. No. 55;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 195 NEALEK 200

RESULT 10

A21762 neurofilament triplet L protein - rat (fragment)

C/Species: *Rattus norvegicus* (Norway rat)

C>Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 13-Aug-1999

C/Accession: A21762

R/Julien, J.P.; Ramchandran, K.; Grosfeld, F.

Biochim. Biophys. Acta 825, 398-404, 1985

A>Title: Cloning of a cDNA encoding the smallest neurofilament protein from the rat.

A/Reference number: A21762; MID:85252830; PMID:3925999

A/Accession: A21762

A>Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-287 <JUL>

A/Cross-references: GB:M25638; NID:9205683; PIDN:AAA1694.1; PID:9205684

C/Superfamily: cytoskeletal keratin

C:Keywords: coiled coil

Query Match 100.0%; Score 29; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 130 NEALEK 135

RESULT 11

AD1614

Acetyl-CoA acetyltransferase homolog lin1453 [imported] - *Listeria innocua* (strain Clp1)

C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AD1614

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baghero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dusurget, O.; Entlian, K.D.; Fsihl, H.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maltournam, A.; Ma

OK, C.; Schueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,

A:Title: Comparative genomics of *Listeria* species

A:Reference number: AB1077; MUID:21537279; PMID:11679666

A:Accession: AD1614

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-389 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC96684.1; PID:g16413926; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lin1453

C:Superfamily: acetyl-CoA acetyltransferase

Query Match 100.0%; Score 29; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 294 NEALEK 299

RESULT 12

A86814

phosphoribosylamine-glycine ligase [imported] - *Lactococcus lactis* subsp. *lactis* (strain

C:Species: *Lactococcus lactis* subsp. *lactis*

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: A86814

R:Bojoltin, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ss

A:Reference number: A86825; MUID:21235186; PMID:11337471

A:Accession: A86814

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-613 <STO>

A:Cross-references: GB:AE005176; PID:g12724511; PIDN:AAK05611.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: purD

C:Superfamily: phosphoribosylamine-glycine ligase; phosphoribosylamine-glycine ligase hc

Query Match 100.0%; Score 29; DB 2; Length 413;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 240 NEALEK 245

RESULT 13

AH1030

Probable exported protein STY4570 [imported] - *Salmonella enterica* subsp. *enterica*

C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh

A:Note: this species has also been called *Salmonella typhi*

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001

C:Accession: AH1030

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Chur

th, T.; Connerton, P.; Cronin, A.; Davys, P.; Davies, R.M.; Dowd, L.; White, N.; Pa

, S.; Moule, S.; O'Garra, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens

A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica*

A:Reference number: AB0502; PMID:11677608

A:Accession: AH1030

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-492 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD09345.1; PID:g16505345; GSPDB:GN00176

C:Genetics:

A:Gene: STY4570

Query Match 100.0%; Score 29; DB 2; Length 492;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 121 NEALEK 126

RESULT 14

QEMSL

neurofilament triplet L protein - mouse

N:Alternate names: 68k neurofilament protein; NF-L(low) protein; type IV IF protein

C:Species: *Mus musculus* (house mouse)

C:Date: 31-Mar-1988 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999

C:Accession: A25227; A26562; A43772; A41012; I55316

R:Lewis, S.A.; Cowan, N.J.

Mol. Cell. Biol. 6, 1529-1534, 1986

A:Title: Anomalous placement of introns in a member of the intermediate filament mu

A:Reference number: A25227; MUID:87064433; PMID:3785173

A:Accession: A25227

A:Molecule type: DNA

A:Residues: 1-543 <LEM>

A:Cross-references: GB:M13016; NID:g200023; PIDN:AA39810.1; PID:g387492

A:Note: the authors translated the codon GGC for residue 5 as Ala, ACA for residue 1

as Glu

R:Lewis, S.A.; Cowan, N.J.

J. Cell Biol. 100, 843-850, 1985

A:Title: Genetics, evolution, and expression of the 68,000-mol-wt neurofilament pro

A:Reference number: A26562; MUID:85131334; PMID:3919033

A:Accession: A26562

A:Molecule type: mRNA

A:Residues: 242-543 <LE2>

A:Cross-references: GB:X02165

A:Experimental source: Brain

R:Julien, J.P.; Meyer, D.; Flavell, D.; Hurst, J.; Grosved, F.

Brain Res. Mol. Brain Res. 1, 243-250, 1986

A:Title: Cloning and developmental expression of the murine neurofilament gene fam1

A:Reference number: A43772

A:Accession: A43772

A:Molecule type: mRNA

A:Residues: 1-5, 'Y', 7-8, 'Y', 10-64, 'W', 66-72, 'L', 74-98, 'D', 100-194, 'R', 196-202, 204-2

A:Cross-references: GB:M20480; NID:g200037; PIDN:AA39814.1; PID:g200038

A:Note: the authors translated the codon CGC for residue 195 as Ala

R:Sligh, R.R.; Nixon, R.A.

J. Biol. Chem. 266, 18861-18867, 1991

A:Title: Identification of Ser-35 as a major protein kinase A phosphorylation site

A:Reference number: A41012; MUID:92011653; PMID:1717455

A:Accession: A41012

A:Molecule type: protein

A:Residues: 52-57 <SH>

R:Nakahira, K Ikenaka, K.; Mada, K.; Tamura, T.
J. Biol. Chem. 265, 19786-19791, 1990
A>Title: Structure of the 66-kDa neurofilament gene and regulation of its expression.
A:Reference number: 155316; MUID:91060592; PMID:2262621
A:Accession: 155316
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5,'1','7'-8,'Y','10'-28 <RES>
A:Cross-references: GB:M55423; NID:g200027; PIND:AAA39812.1; PID:g554245
A:Comment: This is the most abundant of the three neurofilament proteins and, as the other
C:Genetics:
A:Introns: 349/3: 391/2: 498/1
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil; intermediate filament
F:1-2/2/Domain: head <HEAD>
F:94-123/Domain: coil 1a, alpha-helical rod #status predicted <R1a>
F:126-138/Region: linker 1
F:139-234/Domain: coil 1b, alpha-helical rod #status predicted <R1b>
F:235-256/Region: linker 2
F:257-272/Domain: coil 2a, alpha-helical rod #status predicted <R2a>
F:273-281/Region: linker 2
F:282-401/Domain: coil 2b, alpha-helical rod #status predicted <R2b>
F:404-543/Domain: tail <TAI>
F:404-444/Region: tail subdomain a
F:445-543/Region: tail subdomain b

Query Match	100.0%	Score 29;	DB 1;	Length 543;
Best Local Similarity	100.0%	Pred. No. 1.2e+02;		
Matches	6;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	1	NEALEK	6
Db	327	NEALEK	332

RESULT 15

neurofilament triplet L protein - human
N/Alternate names: neurofilament light polypeptide (68k)
N/Contents: Glu-50 brain peptide
C/Species: Homo sapiens (man)
C/Date: 29-Jan-1993 #sequence, revision 29-Jan-1993 #text, change 21-Jul-2000
C/Accession: S07144; 152832; A60703
R/Julien, J.P.; Grosfeld, F.; Yazdanbakhsh, K.; Flavell, D.; Meljzer, D.; Mushynski, W.
Biochim. Biophys. Acta 909, 10-20, 1987.
A>Title: The structure of a human neurofilament gene (NF-L): a unique exon-intron organization
A/Reference number: S07144; MUID:8721213; PMID:3034332
A/Accession: S07144
A/Molecule type: DNA
A/Residues: 1-544 <TUL>
A/Cross-references: EMBL:X05608; NID:q1495072; PIDN:CA929097.1; PID:q1279504
A/Note: the authors translated the codon ATG for residue 366 as Asn
R/Pospelov, V.A.; Pospelova, T.V.; Julien, J.P.
Cell Growth Differ. 5, 187-196, 1994
A>Title: Ap-1 and Krox-24 transcription factors activate the neurofilament light gene promoter
A/Reference number: 152832; MUID:94235564; PMID:8180132
A/Accession: 152832
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-10 <POS>
A/Cross-references: GB:S70309; NID:q547176; PIDN:AMD14057.1; PID:q4261757
R/Nomata, Y.; Watanabe, T.; Wada, H.
J. Biochem. 93, 825-831, 1983
A>Title: Highly acidic proteins from human brain: purification and properties of Glu-50
A/Reference number: A60703; MUID:83265667; PMID:6135695
A/Accession: A60703
A/Molecule type: protein
A/Residues: 469-472, 'D', 474 <NOM>
A/Experimental source: Glu-50 brain peptide
A/Note: this acidic protein is named for its greater than fifty per cent glutamic acid (Glu) content
C/genetics:
A/Gene: GDB:NEFL, NFL
A/Cross-references: GDB:120227; OMIM:162280

A:Map position: 8p21-8p21
A:Introns: 349/3; 391/2; 498/1
C:Superfamily: cytoskeletal keratin
C:Keywords: brain; coiled coil; intermediate filament
F:469-544/Product: Glu-50 peptide #status predicted <E50>

Query Match	100.0%;	Score 29;	DB 2;	Length 544;
Best Local Similarity	100.0%;	Pred. No. 1.2e+02;		
Matches	6;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

QY	1	NEALEK	6
Db	327	NEALEK	332

Search completed: November 13, 2002, 13:23:24
Job time : 10.462 secs

```

-1.rpr                                     Page 5

A:Map position: 8p21-8p21
A:Introns: 349/3; 391/2; 498/1
C:Superfamily: cytoskeletal keratin
C:Keywords: brain; coiled coil; intermediate filament
F:469-544/Product: Glu-50 peptide #status predicted <E50>

Query Match      100.0%; Score 29; DB 2; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NEALEK 6
Db      327 NEALEK 332

Search completed: November 13, 2002, 13:23:24
Job time : 10.462 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 4.97872 Seconds
(without alignments)
49.984 Million cell updates/sec

Title: US-09-856-086-1
Perfect score: 29
Sequence: 1 NEALEK 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	180	1 LACB_SHEEP	P02757 ovls aries
2	29	100.0	183	1 IF3_SERMA	P33320 serralia ma
3	29	100.0	183	1 IF3_YERPE	Q82d46 yerstia pe
4	29	100.0	185	1 ATPF_CYAPA	P48084 cyanophora
5	29	100.0	207	1 HIS2_CAMTE	Q9pm71 campylobact
6	29	100.0	412	1 PUR2_LACLA	Q92t44 lactococcus
7	29	100.0	539	1 MY53_HYDAT	P39922 hydra atten
8	29	100.0	541	1 NFL_RAT	P19527 rattus norv
9	29	100.0	542	1 NFL_MOUSE	P08551 mus musculi
10	29	100.0	543	1 NFL_HUMAN	P07196 homo sapien
11	29	100.0	548	1 NFL_PTG	P02547 sus scrofa
12	29	100.0	554	1 NFL_BOVIN	P02348 bos taurus
13	29	100.0	555	1 NFL_COTJA	Q02316 coturnix co
14	29	100.0	1829	1 DPOL_THEST	Q33845 thermococcu
15	29	93.1	100	1 Y050_MYCPU	Q98tf9 mycoplasma
16	27	93.1	143	1 PTMA_STAAW	P17875 staphylococ
17	27	93.1	143	1 PTMA_STACA	P17876 staphylococ
18	27	93.1	366	1 YF59_STAAW	Q99tc7 staphylococ
19	27	93.1	372	1 ENO_CHLRE	P31683 chlamydomon
20	27	93.1	481	1 ATPB_STIAU	P42469 stigmatella
21	27	93.1	886	1 LEUR_YEAST	P08638 saccharomyc
22	26	89.7	138	1 SODM_MYCCE	P53443 mycobacteri
23	26	89.7	141	1 HBAD_BRACH	P04240 branta cana
24	26	89.7	158	1 CLT4_BOVIN	Q9xsa7 bos taurus
25	26	89.7	186	1 YCEB_SALTY	P40822 salmoneilla
26	26	89.7	200	1 COAE_LISIN	Q92bf2 listeria in
27	26	89.7	220	1 COAE_LISMO	Q8y6w8 listeria mo
28	26	89.7	200	1 EXB8_NEIMC	P95375 neisseria m
29	26	89.7	230	1 YX09_CAEEL	O11115 caenorhabdi
30	26	89.7	253	1 CLT4_HUMAN	Q9y696 homo sapien
31	26	89.7	253	1 CLT4_RAT	Q92bw7 rattus norv
32	26	89.7	268	1 HISJ_NEIGO	P06158 neisseria g
33	26	89.7	275	1 FABI_HELPY	Q9znm7 helicobacte

34	26	89.7	275	1 FABI_HELPY	O24990 helicobacte
35	26	89.7	285	1 GS39_BACSU	P80873 bacillus su
36	26	89.7	290	1 RPOD_AERPE	O9y653 aeropyrum p
37	26	89.7	358	1 Y269_MYCPN	P75395 mycoplasma
38	26	89.7	367	1 CADH_MAIZE	O24562 zea mays (m
39	26	89.7	383	1 TRIS_STACH	O59947 stachydicty
40	26	89.7	398	1 TAP_DROME	O16867 drosophila
41	26	89.7	420	1 PUR2_LISIN	Q928p4 listeria in
42	26	89.7	488	1 MSRA_STAAP	P23212 staphylococ
43	26	89.7	566	1 Y397_MYCCE	P47637 mycoplasma
44	26	89.7	656	1 ACDV_MOUSE	P50544 mus musculi
45	26	89.7	798	1 V120_HSV6G	P30003 human herpe

ALIGNMENTS

RESULT 1	ID	LACB_SHEEP	STANDARD	PRT	180 AA.
AC	P02757				
DT	21-JUL-1986 (Rel. 01, Created)				
DT	01-AUG-1988 (Rel. 08, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	Beta-lactoglobulin 1/B, 2/A, and 3/C precursor.				
OS	Ovis aries (Sheep), and				
OS	Ovis orientalis musimon (Mouflon).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Caprinae; Ovis.				
OX	NCBI_Taxid=9940, 9938;				
RN	[1]				
RP	SEQUENCE FROM N.A. (BLG 1 AND 2).				
RC	SPECIES-SHEEP;				
RX	MEDLINE=88172489; PubMed=3351935;				
RA	All S., Clark A.J.;				
RT	"Characterization of the gene encoding ovine beta-lactoglobulin.				
RT	Similarity to the genes for retinol binding protein and other				
RT	secretory proteins."				
RL	J. Mol. Biol. 199;415-426(1988).				
RN	[2]				
RP	SEQUENCE FROM N.A. (BLG 1).				
RC	SPECIES-SHEEP;				
RX	MEDLINE=87049827; PubMed=3096387;				
RA	Gaye P., Hue-delahate D., Mercier J.-C., Soulier S., Villette J.-L.,				
RT	Furet J.-P.;				
RT	"Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and mRNA				
RT	levels during functional differentiation of the mammary gland."				
RL	Biochimie 68:1097-1107(1986).				
RN	[3]				
RP	SEQUENCE FROM N.A. (BLG 1).				
RC	SPECIES-SHEEP;				
RX	MEDLINE=89057492; PubMed=3194215;				
RA	Harris S., All S., Anderson S., Archibald A.L., Clark A.J.;				
RT	"Complete nucleotide sequence of the genomic ovine beta-lactoglobulin				
RT	gene."				
RL	Nucleic Acids Res. 16:10379-10380(1988).				
RN	[4]				
RP	SEQUENCE FROM N.A. (BLG 1 AND 2).				
RC	SPECIES-SHEEP;				
RX	MEDLINE=91007276; PubMed=1976573;				
RA	All S., McClenaghan M., Simons J.P., Clark A.J.;				
RT	"Characterisation of the alleles encoding ovine beta-lactoglobulins A				
RT	and B."				
RL	Gene 91:201-207(1990).				
RN	[5]				
RP	SEQUENCE OF 19-180 (BLG 2).				
RC	SPECIES-SHEEP;				
RX	MEDLINE=80219294; PubMed=6155985;				
RA	Prieux G., Braunitzer G., Kolde H.-J.;				
RT	"Primary structure of ovine beta-lactoglobulin."				
RL	Arch. Int. Physiol. Biochim. 88:B45-B46(1980).				
RN	[6]				

RP SEQUENCE OF 19-180 (BLG 3).
 RC SPECIES=Sheep;
 RA MEDLINE=89374823; PubMed=2775495;
 RA Erhardt G., Godovac-Zimmermann J., Conti A.;
 RT "Isolation and complete primary sequence of a new ovine wild-type
 RT beta-lactoglobulin C.";
 RT Biol. Chem. Hoppe-Seyler 370:757-762(1989).
 RN [7]
 RP SEQUENCE OF 19-180 (BLG B).
 RC SPECIES=O.o.musimon;
 RA MEDLINE=88106996; PubMed=3426802;
 RA Godovac-Zimmermann J., Conti A., Napolitano L.;
 RT "The complete amino acid sequence of dimeric beta-lactoglobulin from
 RT Biol. Chem. Hoppe-Seyler 368:1313-1319(1987).
 CC -1- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
 CC BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
 CC THAT MOLECULE.
 CC -1- SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS
 CC AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.
 CC -1- MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.
 CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X04520; CAA28204.1; -
 DR EMBL: X12817; CAA31305.1; -
 DR EMBL: X07004; CAA30059.1; ALT_SEQ.
 DR EMBL: X07005; CAA30059.1; JOINED.
 DR EMBL: X07006; CAA30059.1; JOINED.
 DR EMBL: X07007; CAA30059.1; JOINED.
 DR EMBL: X07008; CAA30059.1; JOINED.
 DR EMBL: X07009; CAA30059.1; JOINED.
 DR EMBL: M32236; AAA31510.1; -
 DR EMBL: M32233; AAA31510.1; JOINED.
 DR EMBL: M32232; AAA31510.1; JOINED.
 DR EMBL: M32235; AAA31510.1; JOINED.
 DR PIR: A03221; IGSN.
 DR PIR: S00132; S00132.
 DR PIR: A30011; A30011.
 DR PIR: B30011; B30011.
 DR PIR: A25136; A25136.
 DR PIR: J00748; J00748.
 DR PIR: J00749; J00749.
 DR PIR: S02136; S02136.
 DR PIR: S04955; S04955.
 DR HSSP: P02754; 1BSQ.
 DR InterPro: IPR002345; Lipocalin.
 DR InterPro: IPR000566; Lipocalin_cytfabp.
 DR Pfam: PF00061; Lipocalin.1.
 DR PRINTS: PR00179; LIPOCALIN.
 DR PROSITE: PS00213; LIPOCALIN.1.
 DR MILK: Whey; Retinol-binding; Transport; Signal; Lipocalin.
 KW SIGNAL
 FT CHAIN 1 180 BETA-LACTOGLOBULIN.
 FT DISULFID 84 178
 FT DISULFID 124 137
 FT DISULFID 134 139
 FT VARIANT 38 38 H -> Y (IN LACTOGLOBULIN 2-A AND 3-C).
 FT VARIANT 166 166 R -> Q (IN LACTOGLOBULIN 3-C).
 SQ SEQUENCE 180 AA; 19921 MW; BA8C82E89E757333 CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 180;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NEALEK 6

DB 148 NEALEK 153
 RESULT 2
 IF3_SERMA
 ID IF3_SERMA STANDARD; PRT; 183 AA.
 AC P3320;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Translation initiation factor IF-3.
 GN INFC.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Serratia.
 OX NCBI_Taxid=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 39937;
 RX MEDLINE=94010248; PubMed=8405963;
 RA Liveris D., Schwartz J.J., Geertman R., Schwartz I.;
 RT "Molecular cloning and sequencing of infc, the gene encoding
 RT translation initiation factor IF3, from four enterobacterial
 RT species.";
 RL FEBS Microbiol. Lett. 112:211-216(1993).
 CC -1- FUNCTION: IF-3 BINDS TO THE 30S RIBOSOMAL SUBUNIT AND SHIFTS THE
 CC EQUILIBRIUM BETWEEN 70S RIBOSOMES AND THEIR 50S AND 30S SUBUNITS IN
 CC FAVOR OF THE FREE SUBUNITS, THUS ENHANCING THE AVAILABILITY OF 30S
 CC SUBUNITS ON WHICH PROTEIN SYNTHESIS INITIATION BEGINS.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE IF-3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L11256; AAC36813.1; -
 DR HSSP: P02999; 2IFE.
 DR InterPro: IPR001288; IF3.
 DR Pfam: PF00707; IF3.1.
 DR ProDom: PD002880; IF3.1.
 DR TIGRFAMs: TIGR00168; Infc.1.
 DR PROSITE: PS00938; IF3.1.
 KW Initiation factor; Protein biosynthesis.
 FT SITE 107 IMPORTANT FOR 30S BINDING
 FT SITE 107 (BY SIMILARITY).
 FT SITE 110 IMPORTANT FOR 30S BINDING
 FT SITE 110 (BY SIMILARITY).
 SQ SEQUENCE 183 AA; 20877 MW; BA13F0F6287603 CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 183;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NEALEK 6
 DB 40 NEALEK 45
 RESULT 3
 IF3_YERPE
 ID IF3_YERPE STANDARD; PRT; 183 AA.
 AC O82DW6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Translation initiation factor IF-3.

GN INFC OR YPO2432.
 OS Yersinia pestis.
 CC Bacteria: Proteobacteria: gamma subdivision; Enterobacteriaceae;
 CC Yersinia.
 CC NCBI_TaxID=632;
 CC (1)
 CC SEQUENCE FROM N.A.
 CC STRAIN=CO-92 / Blovat Orientalis;
 CC MEDLINE=21470413; PubMed=11586360;
 CC Parikhil J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 CC Partridge M.B., Sedalshia M., James K.D., Churcher C., Mungall K.L.,
 CC Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Parraga A.M.,
 CC Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 CC Leatherell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,
 CC Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.,
 CC "Genome sequence of Yersinia pestis, the causative agent of plague,"
 CC Nature 413:523-527(2001).
 CC -1- FUNCTION: IF-3 binds to the 30S ribosomal subunit and shifts the
 CC equilibrium between 70S ribosomes and their 50S and 30S subunits in
 CC favor of the free subunits, thus enhancing the availability of 30S
 CC subunits on which protein synthesis initiation begins.
 CC -1- SUBUNIT: Monomer (by similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE IF-3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: A014152; CAC91337.1; .
 CC InterPro: IPR001288; IF3.
 CC Pfam: PF00707; IF3; 1.
 CC TRIGRAMS: TIGR00168; Inf3; 1.
 CC PROSITE: PS00938; IF3; 1.
 CC Initiation factor; Protein biosynthesis; Complete proteome.
 CC KW SEQUENCE 183 AA; 20936 MW; F6A924C69D0E858E CRC64;
 CC SQ
 CC
 CC Query Match 100.0%; Score 29; DB 1; Length 183;
 CC Best Local Similarity 100.0%; Pred. No. 11;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 NEALEK 6
 CC Db 40 NEALEK 45
 CC
 CC RESULT 4
 CC ATPE_CYAPA STANDARD; PRT; 185 AA.
 CC AC P48084;
 CC DT 01-FEB-1996 (Rel. 33, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE ATP synthase B chain (EC 3.6.3.14) (Subunit 1).
 CC GN ATPF.
 CC OS Cyanophora paradoxa.
 CC CC Cyanella.
 CC CC Eukaryote: Glaucocystophyceae; Cyanophoraceae; Cyanophora.
 CC NCBI_TaxID=2762;
 CC (1)
 CC SEQUENCE FROM N.A.
 CC STRAIN=LB555 / Pringsheim;
 CC Stirewalt V.L., Michalowski C.B., Loeffelhardt W., Bohner H.J.,
 CC Bryant D.A.;
 CC "Nucleotide sequence of the cyanella DNA from Cyanophora paradoxa,"
 CC Plant Mol. Biol. Rep. 13:327-332(1995).
 CC (2)
 CC SEQUENCE FROM N.A.

RC STRAIN=LB555 / Pringsheim;
 RA Loeffelhardt W., Stirewalt V.L., Michalowski C.B., Annarella M.,
 RA Farley J.Y., Schluchter W.M., Chung S., Newmann-Spallart C.,
 RA Steiner J.M., Jakowitsch J., Bohner H.J., Bryant D.A.;
 RT "The complete sequence of the cyanella genome of Cyanophora paradoxa:
 RT the genetic complexity of a primitive plastid,"
 RT (in) Schenk H.E.A., Herrmann R., Jeon K.W., Mueller N.E.,
 RL Schwemmer W. (eds.);
 RL Eukaryotism and Symbiosis, pp.40-48, Springer-Verlag, Heidelberg
 RL (1997).
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
 CC H(+) (out).
 CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
 CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
 CC CORES - ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
 CC HAS THREE MAIN SUBUNITS: A, B AND C.
 CC -1- SUBCELLULAR LOCATION: CYANELLE THYLAKOID MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE ATPASE B CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U30821; AAA81255.1; .
 CC InterPro: IPR002146; ATPsyn_B/B_sub.
 CC Pfam: PF00430; ATP-synt_B; 1.
 CC Hydrogen ion transport; Transmembrane; CF(0); Cyanella.
 CC KW SEQUENCE 185 AA; 21256 MW; 87BA778381BEB87 CRC64;
 CC SQ
 CC
 CC Query Match 100.0%; Score 29; DB 1; Length 185;
 CC Best Local Similarity 100.0%; Pred. No. 11;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 1 NEALEK 6
 CC Db 77 NEALEK 82
 CC
 CC RESULT 5
 CC HIS2_CAMJE STANDARD; PRT; 207 AA.
 CC ID HIS2_CAMJE
 CC AC 09PM71;
 CC DT 15-JUN-2002 (Rel. 41, Created)
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Histidine biosynthesis bifunctional protein his2 [includes:
 CC DE Phosphoribosyl-AMP cyclohydrolase (EC 3.5.4.19) (PRA-CH);
 CC DE Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31) (PRA-PH)].
 CC GN HIS2 OR HISIE OR CUI604.
 CC OS Campylobacter jejuni.
 CC CC Bacteria: Proteobacteria; epsilon subdivision; Campylobacter group;
 CC CC Campylobacter.
 CC NCBI_TaxID=197;
 CC (1)
 CC SEQUENCE FROM N.A.
 CC STRAIN=NCTC 11168;
 CC MEDLINE=20150912; PubMed=10688204;
 CC Parikhil J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 CC Basham D., Chillingworth T., Davies R.M., Felwell T., Holroyd S.,
 CC Jagsels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 CC Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 CC Whitehead S., Barrett B.G.;
 CC "The genome sequence of the food-borne pathogen Campylobacter jejuni
 CC RT reveals hypervariable sequences,"
 CC Nature 403:665-668(2000).
 CC -1- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-ATP + H(2)O = 1-(5-
 CC phosphoribosyl)-AMP + diphosphate.
 CC -1- CATALYTIC ACTIVITY: 1-(5-phosphoribosyl)-AMP + H(2)O = 1-(5-
 CC phosphoribosyl)-5-(1-(5-

```

CC      phosphoribosyl(methyl)adenine[midazole-4-carboxamide.
CC      -1- PATHWAY: Histidine biosynthesis; second step.
CC      -1- PATHWAY: Histidine biosynthesis; third step.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE PRA-CH
CC      FAMILY.
CC      -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE PRA-PH
CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: AL139079; CAB73592.1;
CC      DR InterPro: IPR002496; PRA-CH.
CC      DR InterPro: IPR002497; PRA-PH.
CC      DR Pfam: PF01502; PRA-CH; 1.
CC      DR Pfam: PF01503; PRA-PH; 1.
CC      DR ProDom: PD002610; PRA-CH; 1.
CC      DR ProDom: PD002611; PRA-PH; 1.
CC      KW Histidine biosynthesis; Multifunctional enzyme; Hydrolase;
CC      Complete proteome.
CC      FT DOMAIN 1 117 PHOSPHORIBOSYL-AMP CYCLOHYDROLASE.
CC      FT DOMAIN 118 207 PHOSPHORIBOSYL-ATP PYROPHOSPHOHYDROLASE.
CC      SQ SEQUENCE 207 AA; 23175 MW; F5AE6374D6E99C1 CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 1; Length 207;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NEALER 6
Db 40 NEALER 45

RESULT 6
PUN2_LACLA STANDARD; PRT; 412 AA.
ID PUN2_LACLA
AC 09ZF44:
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-UN-2002 (Rel. 41, Last annotation update)
DE Phosphoribosylamine-glycine ligase (EC 6.3.4.13) (GARS) (Glycinamide
DE ribonucleotide synthetase) (Phosphoribosylglycinamide synthetase).
GN PURD OR IL1513.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CHCC373;
RA MEDLINE-99013630; PubMed-9797284;
RA Nilsson D., Klistrup M.;
RT "Cloning and expression of the Lactococcus lactis purDEK genes,
RT required for growth in milk."
RL Appl. Environ. Microbiol. 64:4321-4327(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-IL1403;
RA MEDLINE-21235186; PubMed-11337471;
RA Bolotin A., Wincker P., Mauger S., Jallion O., Malarme K.,
RA Welzenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -1- CATALYTIC ACTIVITY: ATP + 5-phospho-D-ribosylamine + glycine -> ADP
CC + phosphate + N(1)-(5-phospho-D-ribosyl)glycinamide.
CC -1- PATHWAY: De novo purine biosynthesis; second step.
CC -1- SIMILARITY: BELONGS TO THE GARS FAMILY.

```

```

CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: AJ000883; CA004374.1;
CC      DR HSPB; AE006382; AAK05611.1; ALT_INT.
CC      DR HSPB; P15640; IGSO.
CC      DR InterPro: IPR000115; Gars.
CC      DR Pfam: PF01071; GARS; 1.
CC      DR Pfam: PF02842; GARS_B; 1.
CC      DR Pfam: PF02843; GARS_C; 1.
CC      DR Pfam: PF02844; GARS_N; 1.
CC      DR TIGRfam: TIGR00877; purD; 1.
CC      DR PROSITE: PS00184; GARS; 1.
CC      KW Purine biosynthesis; Ligase; Complete proteome.
CC      FT CONFLICT 46 46 D -> H (IN REF. 1).
CC      FT CONFLICT 215 215 G -> E (IN REF. 1).
CC      SQ SEQUENCE 412 AA; 44288 MW; EAAD0D6997E1C02 CRC64;

Query Match
Best Local Similarity 100.0%; Score 29; DB 1; Length 412;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 NEALER 6
Db 239 NEALER 244

RESULT 7
MYS3_HYDAT STANDARD; PRT; 539 AA.
ID MYS3_HYDAT
AC P39922;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, clone 203 (Fragment).
DE Hydra attenuata (Hydra) (Hydra vulgaris).
OS Eukaryota; Metazoa; Chordata; Hydrozoa; Hydrozoa; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakano M.Y., Stidwell R.P.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SIMILARITY: BELONGS TO THE MYOSIN HEAVY CHAIN FAMILY. STRONGEST,
CC TO OTHER NON MUSCLE MYOSINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: L35595; AAA29216.1;
CC      DR HSPB; P24733; IMPC.
CC      DR InterPro: IPR002928; Myosin_tail.
CC      DR Pfam: PF01576; Myosin_tail; 1.
CC      KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
CC      ATP-binding.
CC      FT NON_TER 1 1 GLOBULAR HEAD.
CC      FT DOMAIN 39 >539 RODLIKE TAIL.
CC      FT DOMAIN 39 520 COILED COIL (POTENTIAL).
CC      FT NON_TER 539 539

```

SO SEQUENCE 539 AA; 62319 MW; 9C5AD5664060939D CRC64;

Query Match. 100.0%; Score 29; DB 1; Length 539;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
Db 123 NEALEK 128

RESULT 8

NEFL_RAT STANDARD; PRT; 541 AA.
ID NEFL_RAT
AC P19527; Q63367;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
GN NEFL OR NEFL OR NF68.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90184052; PubMed=2516804;
RA Chin S.S., Liem R.K.H.;
RT "Expression of rat neurofilament proteins NF-L and NF-M in
transfected non-neuronal cells.";
RL Eur. J. Cell Biol. 50:475-490(1989).
RN [2]
RP SEQUENCE OF 197-483 FROM N.A.
RX MEDLINE=85252830; PubMed=3925999;
RA Julien J.-P., Ramchandran K., Grosved F.;
RT Cloning of a cDNA encoding the smallest neurofilament protein from
the rat.";
RL Biochim. Biophys. Acta 825:398-404(1985).
RN [3]
RP SEQUENCE OF 1-10 FROM N.A.
RX MEDLINE=95264348; PubMed=7745611;
RA Reeben M., Neuman T., Palgi J., Palm K., Paalme V., Saarma M.;
RT Characterization of the rat light neurofilament (NF-L) gene promoter
and identification of NGF and cAMP responsive regions.";
RL J. Neurosci. Res. 40:177-188(1995).
RN [4]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=93346421; PubMed=8344946;
RA Dong D.L.-Y., Xu Z.-S., Chevrier M.R., Cotter R.J., Cleveland D.W.,
Hart G.W.;
RT "Glycosylation of mammalian neurofilaments. Localization of multiple
O-linked N-acetylglucosamine moieties on neurofilament polypeptides
L and M.";
RL J. Biol. Chem. 266:16679-16687(1993).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
OTHER NEURONAL COMPONENTS OR IONS.
CC -1- PTM: O-GLYCOSYLATED; CONTAINS THREE N-ACETYLGLUCOSAMINE SIDE
CHAINS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
NEUROFILAMENT PROTEINS AND AS THE OTHER NONEPITHELIAL
INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AF031880; AB87069.1; -;
DR EMBL: M25638; AAA1694.1; -;
DR EMBL: X53981; CAA37931.1; -;
DR PIR: A21762; A21762.
DR GLCOSUITE: P19527; -;
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
DR Intermediate filament; Coiled coil; Neuron; Glycoprotein.
KW INIT MET 0
FT FT DOMAIN 1 92
FT FT DOMAIN 93 396
FT FT DOMAIN 397 541
FT FT DOMAIN 541 144
FT FT DOMAIN 144 137
FT FT DOMAIN 137 233
FT FT DOMAIN 233 252
FT FT DOMAIN 252 271
FT FT DOMAIN 271 280
FT FT DOMAIN 280 396
FT FT DOMAIN 396 443
FT FT DOMAIN 443 541
FT FT CARBOHYD 20 20
FT FT CARBOHYD 26 26
FT FT SITE 381 391
FT FT CONFLICT 197 202
FT FT CONFLICT 399 399
FT FT CONFLICT 476 476
FT FT CONFLICT 480 483
SQ SEQUENCE 541 AA; 61204 MW; 0D17839AF226918A CRC64;
Query Match 100.0%; Score 29; DB 1; Length 541;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 NEALEK 6
Db 326 NEALEK 331
RESULT 9
NEFL_MOUSE STANDARD; PRT; 542 AA.
ID NEFL_MOUSE
AC P08551;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
GN NEFL OR NEFL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=87064433; PubMed=3785173;
RA Lewis S.A., Cowan N.J.;
RT "Anomalous placement of introns in a member of the intermediate
filament multigene family: an evolutionary conundrum.";
RL Mol. Cell. Biol. 6:1529-1534(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;

RA	MEDLINE-87158637; PubMed-31038656;
RX	Jullien J.-P., Meyer D., Flavell D., Hurst J., Grosfeld F.;
RT	"Cloning and developmental expression of the murine neurofilament
RL	gene family.";
RM	Brain Res. 387:243-250(1986).
RN	[3]
RP	SEQUENCE OF 241-542 FROM N.A.
RC	TISSUE-Brain:
RX	MEDLINE-85131334; PubMed-3919033;
RA	Lewis S.A., Cowan N.J.;
RT	"Genetics, evolution, and expression of the 68,000-mol-wt
RL	neurofilament protein: Isolation of a cloned cDNA probe.";
RN	J. Cell Biol. 100:843-850(1985).
RP	[4]
SE	SEQUENCE OF 1-27 FROM N.A.
RX	MEDLINE-91060592; PubMed-2246261;
RA	Nakahita K., Ikenaka K., Mada K., Tamura T.A., Furutachi T.;
RT	Mikoshiba K.;
RL	"Structure of the 68-kDa neurofilament gene and regulation of its
RT	expression.";
RJ	J. Biol. Chem. 265:19786-19791(1990).
-I-	FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
-I-	AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
-I-	DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGISH
CC	THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC	PROTEINS ARE DUE TO THE TAILPEACE EXTENSIONS. THIS REGION MAY FORM
CC	A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC	OTHER NEURONAL COMPONENTS OR IONS.
CC	-I- MISCELLANEOUS: NP-L IS THE MOST ABUNDANT OF THE THREE
CC	INTERMEDIATE FILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC	FILAMENTS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC	FILAMENTS.
-I-	SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC	or send an email to license@isb-sib.ch).

DR	EMBL: X02165; CAB51616.1; -
DR	EMBL: M20480; AAA39814.1; -
DR	EMBL: M13016; AAA39810.1; -
DR	EMBL: M55423; AAA39812.1; -
DR	PIR: A25227; QFMSL.
DR	MCD: MGI:97313; NFL.
DR	InterPro: IPRO01664; IF.
DR	Pfam: PF00038; filament. 1.
DR	PROSITE: PS00226; IF_1.
KW	Intermediate filament; Coiled coil; Neurone; Glycoprotein.
FT	INIT_MET 0
FT	DOMAIN 1 92 HEAD.
FT	DOMAIN 93 396 ROD.
FT	DOMAIN 397 542 TAIL.
FT	DOMAIN 93 124 COIL 1A.
FT	DOMAIN 125 137 LINKER 1.
FT	DOMAIN 138 233 COIL 1B.
FT	DOMAIN 234 252 LINKER 12.
FT	DOMAIN 272 280 COIL 2A.
FT	DOMAIN 281 396 LINKER 2.
FT	DOMAIN 397 443 COIL 2B.
FT	DOMAIN 444 542 TAIL, SUBDOMAIN A. (ACIDIC).
FT	CARBOND 20 20 TAIL, SUBDOMAIN B (BY SIMILARITY).
FT	CARBOND 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
FT	CARBOND 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
FT	SITE 381 391 EPTPEO (RECOGNIZED BY IF-SPECIFIC MONOCLONAL ANTIBODY).
CONFLICT	5 5 Y -> S (IN REF. 1).
CONFLICT	8 8 Y -> I (IN REF. 1).
CONFLICT	64 64 M -> K (IN REF. 1).
CONFLICT	72 72 V -> L (IN REF. 2).

```

CC FT CONFLICT 98 98 R -> H (IN REF. 1).
CC FT CONFLICT 194 194 D -> A (IN REF. 1).
CC FT CONFLICT 202 202 MISSING (IN REF. 2).
CC FT CONFLICT 239 239 Y -> I (IN REF. 1).
CC SQ SEQUENCE 542 AA; 61448 MW; 8BE9B8C6F083D8C CRC64;
CC
CC Query Match 100.0%; Score 29; DB 1; Length 542;
CC Best Local Similarity 100.0%; Pred. No. 35;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0
CC
CC QY 1 NEALEK 6
CC |||||
CC DB 326 NEALEK 331
CC
CC RESULT 10
CC NFL_HUMAN
CC ID NFL_HUMAN STANDARD; PRT; 543 AA.
CC AC P07196; Q16154;
CC DT 01-APR-1988 (Rel. 07, Created)
CC DT 01-APR-1988 (Rel. 07, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Neurofilament triplet L protein (68 kDa neurofilament protein)
CC DE (Neurofilament light polypeptide) (NF-L).
CC GN NFL OR NFL OR NF68.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CC OX NCBI_Taxid=9606;
CC
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE-87214213; PubMed-3034332;
CC RA Julien J.-P., Grosved F., Yazdankhah K., Flavell D., Meijer D.,
CC RA Mushynski W.;
CC RT "The structure of a human neurofilament gene (NF-L): a unique exon-
CC RT intron organization in the intermediate filament gene family.";
CC RL Blochlm. Biophys. Acta 909:10-20(1987).
CC
CC [2]
CC RN SEQUENCE OF 1-10 FROM N.A.
CC RP MEDLINE-94235564; PubMed-8180132;
CC RX Pospelova V.A., Pospelova T.V., Julien J.-P.;
CC RA "AR-1 and Krox-24 transcription factors activate the neurofilament
CC RT light gene promoter in P19 embryonal carcinoma cells.";
CC RL Cell Growth Differ. 5:187-196(1994).
CC
CC [3]
CC RP VARIANT CMT2E PRO-332.
CC RX MEDLINE-20307176; PubMed-10841809;
CC RA Mersyanova I.V., Perpelova A.V., Polyakov A.V., Stukov V.F.,
CC RA Dadaei E.L., Oparin R.B., Petlin A.N., Evgarfov O.V.;
CC RT "A new variant of Charcot-Marie-Tooth disease type 2 is probably the
CC RT result of a mutation in the neurofilament-light gene.";
CC RL Am. J. Hum. Genet. 67:37-46(2000).
CC
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC
CC -1- DISEASE: DEFECTS IN NEFL ARE A CAUSE OF CHARCOT-MARIE-TOOTH
CC DISEASE TYPE 2E (CMT2E).
CC
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -1- DATABASE: NAME-INHERITED peripheral neuropathies mutation db;
CC WWW="http://molgen-www.uia.ac.be/CMTMutations/".
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.lsb-slb.ch/announce/> or send an email to license@lsb-slb.ch).

CC EMBL; X05608; CAA29097.1; -
 CC EMBL; S70309; AADI4057.1; -
 CC PIR; S07144; S07144.
 CC Genew; HGNC:7739; NEFL.
 CC MIM; 162280; -
 CC InterPro; IPR001664; IF.
 CC Pfam; PF00038; Filament; 1.
 CC PROSITE; PS00226; IF; 1.
 CC Intermediate filament; Coiled coil; Neurone; Glycoprotein;
 CC Disease mutation; Charcot-Marie-Tooth disease.
 CC INIT_MET 0
 CC DOMAIN 1 91 HEAD.
 CC DOMAIN 92 396 ROD.
 CC DOMAIN 397 543 TAIL.
 CC DOMAIN 92 123 COIL_1A.
 CC DOMAIN 124 136 LINKER_1.
 CC DOMAIN 137 234 COIL_1B.
 CC DOMAIN 235 252 LINKER_12.
 CC DOMAIN 253 271 COIL_2A.
 CC DOMAIN 272 280 LINKER_2.
 CC DOMAIN 281 396 COIL_2B.
 CC DOMAIN 397 443 TAIL, SUBDOMAIN A.
 CC DOMAIN 444 543 TAIL, SUBDOMAIN B (ACIDIC).
 CC CARBOHYD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
 CC CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
 CC SITE 381 391 EPTOPE (RECOGNIZED BY IF-SPECIFIC MONOCLONAL ANTIBODY).
 CC VARIANT 332 332 O -> P (IN CMT2E).
 CC FT SEQUENCE 543 AA; 61645 MW; 7A0F1AD5BED226 CRC64;
 CC SQ

Query Match 100.0%; Score 29; DB 1; Length 543;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
 Db 326 NEALEX 331

RESULT 11

NEFL_PIG STANDARD; PRT; 548 AA.
 ID NEFL_PIG
 AC P02547;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-OCT-1996 (Rel. 34, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L).
 DE NEFL.
 GN Sus acrota (Pig).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OC NCBI_TaxID=9823;
 RX MEDLINE=85154583; PubMed=3920075;
 RA Geisler N., Plesmann U., Weber K.;
 RT "The complete amino acid sequence of the major mammalian
 RT neurofilament protein (NF-L)."
 RL FEBS Lett. 182:475-478(1985).
 RN [2]
 RP SEQUENCE OF 1-82 AND 278-548.
 RA Geisler N., Kauffmann E., Fischer S., Plesmann U., Weber K.;
 RT "Neurofilament architecture combines structural principles of
 RT intermediate filaments with carboxy-terminal extensions increasing
 RT in size between triplet proteins."
 RL EXMO J. 2:1295-1302(1983).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,

CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAPFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC PIR; A02963; ORFGL.
 CC InterPro; IPR001664; IF.
 CC Pfam; PF00038; Filament; 1.
 CC PROSITE; PS00226; IF; 1.
 CC Intermediate filament; Coiled coil; Neurone; Glycoprotein.

CC DOMAIN 1 91 HEAD.
 CC DOMAIN 92 395 ROD.
 CC DOMAIN 396 548 TAIL.
 CC DOMAIN 92 123 COIL_1A.
 CC DOMAIN 124 136 LINKER_1.
 CC DOMAIN 137 232 COIL_1B.
 CC DOMAIN 233 251 LINKER_12.
 CC DOMAIN 252 270 COIL_2A.
 CC DOMAIN 271 279 LINKER_2.
 CC DOMAIN 280 395 COIL_2B.
 CC DOMAIN 396 442 TAIL, SUBDOMAIN A.
 CC DOMAIN 443 548 TAIL, SUBDOMAIN B (ACIDIC).
 CC CARBOHYD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
 CC CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
 CC SITE 380 390 EPTOPE (RECOGNIZED BY IF-SPECIFIC MONOCLONAL ANTIBODY).
 CC FT SEQUENCE 548 AA; 61940 MW; 83044813637AC729 CRC64;
 CC SQ

Query Match 100.0%; Score 29; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
 Db 325 NEALEX 330

RESULT 12

NEFL_BOVIN STANDARD; PRT; 554 AA.
 ID NEFL_BOVIN
 AC P02546; P79127;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L) (Micro glutamic acid-rich
 DE protein).
 DE NEFL.
 GN Bos taurus (Bovine).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RX SEQUENCE FROM N.A.
 RC SPRAIN-Holstein; TISSUE=Brain;
 RA Hill W.D., Zhang L., Balin B.J., Sprinkle T.J.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 468-554.
 RA MEDLINE=85154567; PubMed=384373;
 RA Isobe T., Okuyama T.;
 RT "Brain micro glutamic acid-rich protein is the C-terminal endpiece of
 RT the neurofilament 68-kDa protein as determined by the primary
 RT sequence."
 RL FEBS Lett. 182:389-392(1985).
 CC


```

CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U83919; AAB41543.1; -.
CC PIR: A02964; QF80.
CC InterPro: IPR001664; IF.
CC Pfam: PF00038; filament; 1.
CC PROSITE: PS00226; IF; 1.
CC
CC Intermediate filament; Coiled coil; Neutrone.
CC INT_MET 0
CC BY SIMILARITY.
CC FT DOMAIN 1 92 HEAD (BY SIMILARITY).
CC FT DOMAIN 93 396 ROD (BY SIMILARITY).
CC FT DOMAIN 397 554 TAIL (BY SIMILARITY).
CC FT DOMAIN 93 124 COIL 1A.
CC FT DOMAIN 125 137 LINKER 1.
CC FT DOMAIN 138 233 COIL 1B.
CC FT DOMAIN 234 252 LINKER 12.
CC FT DOMAIN 253 271 COIL 2A.
CC FT DOMAIN 272 280 LINKER 2.
CC FT DOMAIN 281 396 COIL 2B.
CC FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
CC FT DOMAIN 444 554 TAIL, SUBDOMAIN B (ACIDIC).
CC FT CONFLICT 494 500 MISSING (IN REF. 2).
CC FT CONFLICT 509 509 A -> AEA (IN REF. 2).
CC SQ SEQUENCE 554 AA; 62514 MW; D772B81CA2C31C1A CnC64;

```

Query Match 100.0%; Score 29; DB 1; Length 554;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
 DB 326 NEALEX 331

RESULT 13
 ID NFL_COTJA STANDARD: PRT: 555 AA.
 AC 002916;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (Neurofilament light polypeptide) (NF-L).
 GN NEFL.
 OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Coturnix.
 OX NCBI_TaxID-93934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE-93224534; PubMed-8468353;
 RA Ohara O., Gehara Y., Miyake T., Teroka H., Kitamura T.;

```

RT "Neurofilament deficiency in quail caused by nonsense mutation in
RT neurofilament-L gene."
RL J. Cell Biol. 121:387-395(1993).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D13223; BAA02504.1; -.
CC DR EMBL: D13222; BAA02503.1; ALT_TERM.
CC DR InterPro: IPR001664; IF.
CC Pfam: PF00038; filament; 1.
CC PROSITE: PS00226; IF; 1.
CC
CC Intermediate filament; Coiled coil; Neutrone.
CC INT_MET 0
CC BY SIMILARITY.
CC FT DOMAIN 1 93 HEAD (BY SIMILARITY).
CC FT DOMAIN 94 397 ROD (BY SIMILARITY).
CC FT DOMAIN 398 555 TAIL (BY SIMILARITY).
CC FT DOMAIN 94 125 COIL 1A.
CC FT DOMAIN 126 138 LINKER 1.
CC FT DOMAIN 139 234 COIL 1B.
CC FT DOMAIN 235 253 LINKER 12.
CC FT DOMAIN 254 272 COIL 2A.
CC FT DOMAIN 273 397 LINKER 2.
CC FT DOMAIN 398 444 TAIL, SUBDOMAIN A.
CC FT DOMAIN 445 555 TAIL, SUBDOMAIN B (ACIDIC).
CC SQ SEQUENCE 555 AA; 62282 MW; 9B57ABDBEBA7712 CnC64;

```

Query Match 100.0%; Score 29; DB 1; Length 555;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEALEX 6
 DB 327 NEALEX 332

RESULT 14
 ID DPOL_THEST STANDARD: PRT: 1829 AA.
 AC 033845;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7).
 GN POL.
 OS Thermococcus sp. (strain TY).
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Thermococcus.
 OX NCBI_TaxID-110163;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE-98094267; PubMed-9434178;
 RA Niehaus F., Frey B., Antiklian G.;
 "Cloning and characterisation of a thermostable alpha-DNA polymerase

```

RT from the hyperthermophilic archaeon Thermococcus sp. TY.
RL Gene 204:153-158(1997).
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
CC + (DNA)(N).
CC -1- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSCRIPTIONAL EXCISION OF THE THREE INTERVENING REGION
CC (INTERINS) FOLLOWED BY PEPTIDE LIGATION.
CC -1- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y13030; CAW73475.1; -.
DR HSSP: P56689; ITGO.
DR InterPro: IPR002064; DNA_pol_B.
DR InterPro: IPR003586; Hedgehog_hlntc.
DR InterPro: IPR003587; Hedgehog_hlntc.
DR InterPro: IPR002203; Intein.
DR InterPro: IPR004042; Intein_endonuc.
DR InterPro: IPR004578; Pol2.
DR Pfam: PF00136; DNA_pol_B_4.
DR Pfam: PF03104; DNA_pol_B_exo; 1.
DR PRINTS: PR00379; INTEIN.
DR SMART: SM00305; Hlntc; 3.
DR SMART: SM00306; Hlntc; 3.
DR SMART: SM00486; POLBc; 1.
DR TIGRFS: TIGR00592; Pol2; 2.
DR PROSITE: PS00116; DNA_POLYMERASE_B; FALSE_NEG.
DR PROSITE: PS50818; INTEIN_C_TER; 3.
DR PROSITE: PS50819; INTEIN_ENDONUCLEASE; 2.
DR PROSITE: PS50817; INTEIN_N_TER; 3.
DR Transferrase: DNA-directed DNA polymerase; DNA replication;
DR DNA-binding; Hydrolase; Endonuclease; Autocatalytic cleavage;
DR Protein splicing.
DR CHAIN 1 409 DNA POLYMERASE, 1ST PART (POTENTIAL).
FT CHAIN 410 769 INTEIN I.
FT CHAIN 770 855 DNA POLYMERASE, 2ND PART (POTENTIAL).
FT CHAIN 856 1392 INTEIN II.
FT CHAIN 1393 1441 DNA POLYMERASE, 3RD PART (POTENTIAL).
FT CHAIN 1442 1598 INTEIN III.
FT CHAIN 1599 1829 DNA POLYMERASE, 4TH PART (POTENTIAL).
SQ SEQUENCE 1829 AA; 211875 MW; A113A8BC57EB9CB3 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 1829;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEX 6
DB 712 NEALEX 717

```

```

RA Chandaud I., Hellig R., Ferris S., Barbe V., Samson D., Gollison F.,
RA Moszer I., Dypvig K., Wroblewski H., Viari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis."; 29:2145-2153(2001).
RL Nucleic Acids Res. 29:2145-2153(2001).
CC -1- SIMILARITY: BELONGS TO THE UPF0133 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL445563; CAC13223.1; -.
DR MyPulst: MYP0_0500; -.
DR InterPro: IPR004401; Cons_hypoth103.
DR InterPro: IPR003727; DUF149.
DR Pfam: PF02575; DUF149; 1.
DR TIGRFS: TIGR00103; Cons_hypoth103; 1.
DR Hypothetical protein; Complete proteome.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 100 AA; 11656 MW; 18A2B651EB40680 CRC64;

Query Match 93.1%; Score 27; DB 1; Length 100;
Best Local Similarity 83.3%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEX 6
DB 76 NEALEX 81

```

Search completed: November 13, 2002, 13:16:49
 Job time : 5.97872 secs

```

RESULT 15
Y050_MYCPU
ID Y050_MYCPU STANDARD; PRT; 100 AA.
AC 098RP9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein MYP0_0500.
GN MYP0_0500.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 19.2766 seconds
(without alignments)
64.134 Million cell updates/sec

Title: US-09-856-086-1

Perfect score: 29
Sequence: 1 NEALEX 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacterioph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	79	16	Q8RB86 thermocane
2	29	100.0	128	15	Q9QFL4 human immun
3	29	100.0	130	15	Q9QFL1 human immun
4	29	100.0	131	15	Q36943 human immun
5	29	100.0	132	15	Q9QFL1 human immun
6	29	100.0	132	15	Q9QFL0 human immun
7	29	100.0	132	15	Q9QFL8 human immun
8	29	100.0	132	15	Q9QFL8 human immun
9	29	100.0	132	15	Q9QFL8 human immun
10	29	100.0	133	15	Q9QFL5 human immun
11	29	100.0	133	15	Q9QFL9 human immun
12	29	100.0	133	15	Q9QFL8 human immun
13	29	100.0	133	15	Q9QFL7 human immun
14	29	100.0	133	15	Q9QFL5 human immun
15	29	100.0	133	15	Q9QFL4 human immun
16	29	100.0	133	15	Q9QFL4 human immun

17	29	100.0	133	15	Q9QFL3 human immun
18	29	100.0	133	15	Q9QFL2 human immun
19	29	100.0	133	15	Q9QFL1 human immun
20	29	100.0	133	15	Q9QFL0 human immun
21	29	100.0	133	15	Q9QFL9 human immun
22	29	100.0	133	15	Q9QFL8 human immun
23	29	100.0	133	15	Q9QFL7 human immun
24	29	100.0	133	15	Q9QFL6 human immun
25	29	100.0	133	15	Q9QFL5 human immun
26	29	100.0	133	15	Q9QFL4 human immun
27	29	100.0	133	15	Q9QFL3 human immun
28	29	100.0	133	15	Q9QFL2 human immun
29	29	100.0	133	15	Q9QFL1 human immun
30	29	100.0	133	15	Q9QFL0 human immun
31	29	100.0	133	15	Q9QFL9 human immun
32	29	100.0	133	15	Q9QFL8 human immun
33	29	100.0	133	15	Q9QFL7 human immun
34	29	100.0	133	15	Q9QFL6 human immun
35	29	100.0	133	15	Q9QFL5 human immun
36	29	100.0	133	15	Q9QFL4 human immun
37	29	100.0	133	15	Q9QFL3 human immun
38	29	100.0	133	15	Q9QFL2 human immun
39	29	100.0	133	15	Q9QFL1 human immun
40	29	100.0	133	15	Q9QFL0 human immun
41	29	100.0	133	15	Q9QFL9 human immun
42	29	100.0	133	15	Q9QFL8 human immun
43	29	100.0	133	15	Q9QFL7 human immun
44	29	100.0	133	15	Q9QFL6 human immun
45	29	100.0	133	15	Q9QFL5 human immun

ALIGNMENTS

RESULT 1
Q8RB86 PRELIMINARY; PRT; 79 AA.
AC Q8RB86;
DT 01-JUN-2002 (TREMREL. 21, Created)
DT 01-JUN-2002 (TREMREL. 21, Last sequence update)
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)
DE Hypothetical protein TTE0936.
GN TTE0936.
GN Thermotoga thermophilus tengcongensis.
OS Thermotoga thermophilus tengcongensis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Thermotoga thermophilus; Thermotogaceae; Thermotogales.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB47 / JCM11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Tan H., Chen R., Wang J., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RT "A complete sequence of T. tengcongensis genome."
RL Genome Res. 12:689-700(2002).
DR EMBL, AE013059; AAM24192.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 79 AA; 9078 MW; B3FC525B10E95FF9 CRC64;
Query Match 100.0%; Score 29; DB 16; Length 79;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NEALEX 6
DB 22 NEALEX 27
RESULT 2
Q9QFL4 PRELIMINARY; PRT; 128 AA.
AC Q9QFL4;
ID Q9QFL4;

DB 72 NEALEK 77

RESULT 4

ID 036943 PRELIMINARY; PRT; 131 AA.

AC 036943;

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)

DE GAG polypeptide (Fragment).

GN GAG.

OS Human immunodeficiency virus type 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RA Bell J.E., Simmonds P., Hughes E.S., Brettle R.P., Morris A., Marsden M., Halcrow K., Hughes E.S., Brettle R.P., "Mosaic structure of the human immunodeficiency virus type 1 genome infecting lymphoid cells and the brain: evidence for frequent in vivo recombination events in the evolution of regional populations." J. Virol. 73:8720-8731(1999).

RL EMBL: AF174984; AAF00263.1; -

DR InterPro: IPR000071; Retrovir_p17.

DR Pfam: PF00540; Gag_p17; 1.

DR PRINTS: PR00234; HIVMATR1X.

KW AIDS; Core protein; Polypeptide.

FT NON_TER 1

FT NON_TER 1

SQ SEQUENCE 128 AA; 14272 MW; AB8349396175E9AA CRC64;

Query Match 100.0%; Score 29; DB 15; Length 128;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 70 NEALEK 75

RESULT 3

ID 090FL1 PRELIMINARY; PRT; 130 AA.

AC 090FL1;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)

DE GAG polypeptide (Fragment).

GN GAG.

OS Human immunodeficiency virus type 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RA Bell J.E., Simmonds P., Hughes E.S., Brettle R.P., Morris A., Marsden M., Halcrow K., Hughes E.S., Brettle R.P., "Mosaic structure of the human immunodeficiency virus type 1 genome infecting lymphoid cells and the brain: evidence for frequent in vivo recombination events in the evolution of regional populations." J. Virol. 73:8720-8731(1999).

RL EMBL: AF174987; AAF00333.1; -

DR InterPro: IPR000071; Retrovir_p17.

DR Pfam: PF00540; Gag_p17; 1.

DR PRINTS: PR00234; HIVMATR1X.

KW AIDS; Core protein; Polypeptide.

FT NON_TER 1

FT NON_TER 1

SQ SEQUENCE 130 AA; 14542 MW; FBB9EE60AFAADC39 CRC64;

Query Match 100.0%; Score 29; DB 15; Length 130;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||

DB 72 NEALEK 77

RESULT 4

ID 036943 PRELIMINARY; PRT; 131 AA.

AC 036943;

DT 01-JAN-1998 (TREMBLrel. 05, Created)

DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)

DE GAG polypeptide (Fragment).

GN GAG.

OS Human immunodeficiency virus type 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RA Bell J.E., Simmonds P., Hughes E.S., Brettle R.P., Morris A., Marsden M., Halcrow K., Hughes E.S., Brettle R.P., "Mosaic structure of the human immunodeficiency virus type 1 genome infecting lymphoid cells and the brain: evidence for frequent in vivo recombination events in the evolution of regional populations." J. Virol. 73:8720-8731(1999).

RL EMBL: AF174917; AAF00263.1; -

DR InterPro: IPR000071; Retrovir_p17.

DR Pfam: PF00540; Gag_p17; 1.

DR PRINTS: PR00234; HIVMATR1X.

KW AIDS; Core protein; Polypeptide.

FT NON_TER 1

FT NON_TER 1

SQ SEQUENCE 131 AA; 14548 MW; ACBF267B07F566C7 CRC64;

Query Match 100.0%; Score 29; DB 15; Length 131;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
|||||
DB 75 NEALEK 80

RESULT 5

ID 090FT1 PRELIMINARY; PRT; 132 AA.

AC 090FT1;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)

DE GAG polypeptide (Fragment).

GN GAG.

OS Human immunodeficiency virus type 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI_TaxID=11676;

RA Bell J.E., Simmonds P., Hughes E.S., Brettle R.P., Morris A., Marsden M., Halcrow K., Hughes E.S., Brettle R.P., "Mosaic structure of the human immunodeficiency virus type 1 genome infecting lymphoid cells and the brain: evidence for frequent in vivo recombination events in the evolution of regional populations." J. Virol. 73:8720-8731(1999).

RL EMBL: AF174917; AAF00263.1; -

DR InterPro: IPR000071; Retrovir_p17.

DR Pfam: PF00540; Gag_p17; 1.

DR PRINTS: PR00234; HIVMATR1X.

KW AIDS; Core protein; Polypeptide.

FT NON_TER 1

FT NON_TER 1

SQ SEQUENCE 132 AA; 14756 MW; B1EC094AE2D43C39 CRC64;

Query Match 100.0%; Score 29; DB 15; Length 132;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 74 NEALEK 79

RESULT 6

O9QF70 ID O9QF70 PRELIMINARY; PRT; 132 AA.
 AC O9QF70;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentiviridae.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CP4;
 RX MEDLINE-99412391; PubMed-10482626;
 RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,
 Bell J.E., Simmonds P.;
 RT "Mosaic structure of the human immunodeficiency virus type 1 genome
 infecting lymphoid cells and the brain: evidence for frequent in vivo
 recombination events in the evolution of regional populations.";
 RL J. Virol. 73:8720-8731(1999).
 DR EMBL: AF174918; AAF00264.1; -
 DR InterPro: IPR000071; Retrovirl_P17.
 DR Pfam: PF00540; Gag_P17; 1.
 DR PRINTS: PR00234; HIVMATRIX.
 KW AIDS; Core protein; Polyprotein.
 FT NON_TER 1
 FT NON_TER 132
 SQ SEQUENCE 132 AA; 14779 MW; 373764F8ABF3DC40 CRC64;

Query Match 100.0%; Score 29; DB 15; Length 132;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 74 NEALEK 79

RESULT 7

O9QFS9 ID O9QFS9 PRELIMINARY; PRT; 132 AA.
 AC O9QFS9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentiviridae.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CP5;
 RX MEDLINE-99412391; PubMed-10482626;
 RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,
 Bell J.E., Simmonds P.;
 RT "Mosaic structure of the human immunodeficiency virus type 1 genome
 infecting lymphoid cells and the brain: evidence for frequent in vivo
 recombination events in the evolution of regional populations.";
 RL J. Virol. 73:8720-8731(1999).
 DR EMBL: AF174919; AAF00265.1; -
 DR InterPro: IPR000071; Retrovirl_P17.

DR Pfam: PF00540; Gag_P17; 1.
 DR PRINTS: PR00234; HIVMATRIX.
 KW AIDS; Core protein; Polyprotein.
 FT NON_TER 1
 FT NON_TER 132
 SQ SEQUENCE 132 AA; 14730 MW; B1EC095872D9AC26 CRC64;

Query Match 100.0%; Score 29; DB 15; Length 132;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 74 NEALEK 79

RESULT 8

O9QFS8 ID O9QFS8 PRELIMINARY; PRT; 132 AA.
 AC O9QFS8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentiviridae.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CP6;
 RX MEDLINE-99412391; PubMed-10482626;
 RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,
 Bell J.E., Simmonds P.;
 RT "Mosaic structure of the human immunodeficiency virus type 1 genome
 infecting lymphoid cells and the brain: evidence for frequent in vivo
 recombination events in the evolution of regional populations.";
 RL J. Virol. 73:8720-8731(1999).
 DR EMBL: AF174920; AAF00266.1; -
 DR InterPro: IPR000071; Retrovirl_P17.
 DR Pfam: PF00540; Gag_P17; 1.
 DR PRINTS: PR00234; HIVMATRIX.
 KW AIDS; Core protein; Polyprotein.
 FT NON_TER 1
 FT NON_TER 132
 SQ SEQUENCE 132 AA; 14699 MW; 4C5ED1DD5F542988 CRC64;

Query Match 100.0%; Score 29; DB 15; Length 132;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 75 NEALEK 80

RESULT 9

O9QFS5 ID O9QFS5 PRELIMINARY; PRT; 132 AA.
 AC O9QFS5;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentiviridae.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CP9;
 RX MEDLINE-99412391; PubMed-10482626;
 RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,

RA Bell J.E., Slimmons P.;
RT "Mosaic structure of the human immunodeficiency virus type 1 genome
infecting lymphoid cells and the brain: evidence for frequent in vivo
recombination events in the evolution of regional populations.";
RL J. Virol. 73:8720-8731(1999).
DR EMBL; AF174923; AAF00269.1; -
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS; PR00234; HIVMATRIX.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT 132
SQ SEQUENCE 132 AA; 14683 MW; 2045A6B15F5439D8 CRC64;

Query Match 100.0%; Score 29; DB 15; Length 132;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
Db 75 NEALEK 80

RESULT 10
O90G00 PRELIMINARY; PRT; 133 AA.
AC O90G00;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Gag polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BG;
RX MEDLINE=99412391; PubMed=10482626;
RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettle R.P.,
RT "Mosaic structure of the human immunodeficiency virus type 1 genome
infecting lymphoid cells and the brain: evidence for frequent in vivo
recombination events in the evolution of regional populations.";
RL J. Virol. 73:8720-8731(1999).
DR EMBL; AF174848; AAF00863.1; -
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS; PR00234; HIVMATRIX.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT 133
SQ SEQUENCE 133 AA; 14876 MW; C3667671E11E620A CRC64;

Query Match 100.0%; Score 29; DB 15; Length 133;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
Db 75 NEALEK 80

RESULT 11
O90F29 PRELIMINARY; PRT; 133 AA.
AC O90F29;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Gag polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.

OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BG;
RX MEDLINE=99412391; PubMed=10482626;
RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettle R.P.,
RT "Mosaic structure of the human immunodeficiency virus type 1 genome
infecting lymphoid cells and the brain: evidence for frequent in vivo
recombination events in the evolution of regional populations.";
RL J. Virol. 73:8720-8731(1999).
DR EMBL; AF174848; AAF00864.1; -
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS; PR00234; HIVMATRIX.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT 133
SQ SEQUENCE 133 AA; 14914 MW; 6C66767C9A82A65 CRC64;

Query Match 100.0%; Score 29; DB 15; Length 133;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
Db 75 NEALEK 80

RESULT 12
O90F28 PRELIMINARY; PRT; 133 AA.
AC O90F28;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Gag polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LO;
RX MEDLINE=99412391; PubMed=10482626;
RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettle R.P.,
RT "Mosaic structure of the human immunodeficiency virus type 1 genome
infecting lymphoid cells and the brain: evidence for frequent in vivo
recombination events in the evolution of regional populations.";
RL J. Virol. 73:8720-8731(1999).
DR EMBL; AF174850; AAF00865.1; -
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
DR PRINTS; PR00234; HIVMATRIX.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT 133
SQ SEQUENCE 133 AA; 14820 MW; B1EC5EC82C835019 CRC64;

Query Match 100.0%; Score 29; DB 15; Length 133;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
Db 75 NEALEK 80

RESULT 13
O90F27 PRELIMINARY; PRT; 133 AA.

AC 090F27;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Lentivirus.
 OX NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RT;
 RX MEDLINE-99412391; PubMed-10482626;
 RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,
 Bell J.E., Simmonds P.;
 RT "Mosaic structure of the human immunodeficiency virus type 1 genome
 infecting lymphoid cells and the brain: evidence for frequent in vivo
 recombination events in the evolution of regional populations.";
 RL J. Virol. 73:8720-8731(1999).
 DR EMBL; AF174851; AAF00866.1; -;
 DR InterPro: IPR000071; Retrovir_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 DR PRINTS; PR00234; HIVMATRIX.
 KW AIDS; Core protein; Polypotein.
 FT NON_TER 1
 FT SEQUENCE 133 AA; 14792 MW; 9CEC5ED1DF9BEC10 CRC64;
 SQ

Query Match 100.0%; Score 29; DB 15; Length 133;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 75 NEALEK 80

RESULT 14
 ID 090F25 PRELIMINARY; PRT; 133 AA.
 AC 090F25;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Lentivirus.
 OX NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RO;
 RX MEDLINE-99412391; PubMed-10482626;
 RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,
 Bell J.E., Simmonds P.;
 RT "Mosaic structure of the human immunodeficiency virus type 1 genome
 infecting lymphoid cells and the brain: evidence for frequent in vivo
 recombination events in the evolution of regional populations.";
 RL J. Virol. 73:8720-8731(1999).
 DR EMBL; AF174851; AAF00866.1; -;
 DR InterPro: IPR000071; Retrovir_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 DR PRINTS; PR00234; HIVMATRIX.
 KW AIDS; Core protein; Polypotein.
 FT NON_TER 1
 FT SEQUENCE 133 AA; 14813 MW; B1FFFC4CD65F5429 CRC64;
 SQ

Query Match 100.0%; Score 29; DB 15; Length 133;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6

DB 75 NEALEK 80

RESULT 15
 ID 090F24 PRELIMINARY; PRT; 133 AA.
 AC 090F24;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 DE GAG polyprotein (Fragment).
 GN GAG.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Lentivirus.
 OX NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RT;
 RX MEDLINE-99412391; PubMed-10482626;
 RA Morris A., Marsden M., Halcrow K., Hughes E.S., Brettie R.P.,
 Bell J.E., Simmonds P.;
 RT "Mosaic structure of the human immunodeficiency virus type 1 genome
 infecting lymphoid cells and the brain: evidence for frequent in vivo
 recombination events in the evolution of regional populations.";
 RL J. Virol. 73:8720-8731(1999).
 DR EMBL; AF174854; AAF00869.1; -;
 DR InterPro: IPR000071; Retrovir_p17.
 DR Pfam: PF00540; Gag_p17; 1.
 DR PRINTS; PR00234; HIVMATRIX.
 KW AIDS; Core protein; Polypotein.
 FT NON_TER 1
 FT SEQUENCE 133 AA; 14813 MW; B1EC5ED1D5F5429 CRC64;
 SQ

Query Match 100.0%; Score 29; DB 15; Length 133;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NEALEK 6
 DB 75 NEALEK 80

Search completed: November 13, 2002, 13:16:04
 Job time : 20.2766 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 34.7021 seconds
(without alignments)
26.879 Million cell updates/sec

Title: US-09-856-086-2
Perfect score: 36
Sequence: 1 LKRVHEE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneseq/emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	36	100.0	7	21	AA05926
2	36	100.0	441	19	AA20612
3	36	100.0	543	22	AA03466
4	33	91.7	15	23	AA07945
5	33	91.7	319	21	AA04280
6	33	91.7	430	23	AA07947
7	33	91.7	465	19	AA05435
8	33	91.7	465	23	AA07739
9	33	91.7	466	21	AA02963
10	33	91.7	466	21	AA02933

11	33	91.7	466	22	AA06348	Human vimentin. H
12	33	91.7	466	23	AA08769	Human pancreatic t
13	33	91.7	475	22	AA09551	Human protein sequ
14	33	91.7	541	22	AA09462	Human protein sequ
15	33	91.7	545	22	AA02547	Human protein sequ
16	33	91.7	598	23	AA08998	L. gasseri Beta-glu
17	33	91.7	3135	15	AA07747	P. falciparum tran
18	33	91.7	3135	15	AA08223	Plasmodium falcipa
19	32	88.9	6	21	AA05930	Test antigen #6 fo
20	32	88.9	142	18	AA05540	H. pylori ORF 129p
21	32	88.9	147	18	AA05528	H. pylori ORF 13ap
22	32	88.9	400	18	AA05569	H. pylori ORF 069p
23	32	88.9	400	19	AA09830	H. pylori GHPO 219
24	32	88.9	433	19	AA02097	Human glial fibril
25	31	86.1	76	21	AA03553	Human secreted pro
26	31	86.1	85	20	AA03610	Extended human sec
27	31	86.1	188	22	AA04210	Human polypeptide
28	31	86.1	189	22	AA02977	Novel human secret
29	31	86.1	212	22	AA06038	Drosophila melanog
30	31	86.1	278	22	AA09457	Human protein sequ
31	31	86.1	381	22	AA01735	Novel signal trans
32	31	86.1	381	22	AA09386	Human protein sequ
33	31	86.1	397	22	AA08247	Drosophila melanog
34	31	86.1	600	22	AA03949	Human polypeptide
35	31	86.1	600	23	AA08951	Human polypeptide
36	31	86.1	600	23	AA01525	Human RNA metaboli
37	31	86.1	613	21	AA04236	Human ORF ORF2133
38	31	86.1	615	22	AA07965	Human protein SEQ
39	31	86.1	615	22	AA04128	Human polypeptide
40	31	86.1	622	22	AA07867	Human protein SEQ
41	31	86.1	642	22	AA09369	Human protein sequ
42	31	86.1	658	22	AA04032	Human polypeptide
43	31	86.1	753	23	AA05002	Listeria monocytog
44	31	86.1	1037	22	AA09358	Human protein sequ
45	30	83.3	64	22	AA04012	Peptide #7630 enco

ALIGNMENTS

RESULT 1
ID AA05926 standard; peptide; 7 AA.
XX
AC AA05926:
XX
DT 17-OCT-2000 (first entry)
XX
DE Test antigen #2 for spongiform and demyelinating disease diagnosis.
XX
KW Human; cow; myelin; myelin neurofilament; immunogen; antigen;
KW bovine spongiform encephalopathy; BSE; multiple sclerosis;
KW Creutzfeldt-Jacob disease; CJD; demyelinating disease; diagnostic test.
XX
OS Bos taurus.
OS Homo sapiens.
XX
PN WO200031545-A1.
XX
PD 02-JUN-2000.
XX
PF 25-NOV-1999; 99WO-GB03936.
XX
PR 26-NOV-1998; 98GB-0025948.
XX
(UNLO) KING'S COLLEGE.
PI Edinger A;
XX WPI: 2000-400194/34.
XX
PT Diagnosing spongiform or demyelinating disease in vertebrates such as
bovine spongiform encephalopathy and Creutzfeldt-Jacob disease comprises

PT Assaying a biological sample for myelin and/or myelin neurofilament
PT antibodies -
XX
XX Claim 5; Page 2; 16pp; English.
XX
CC The present peptide may be used as a test antigen in a kit for diagnosing
CC spongiform or demyelinating disease in vertebrates, including bovine
CC spongiform encephalopathy (BSE), multiple sclerosis (MS) and
CC Creutzfeldt-Jacob disease (CJD). Peptides used in the kit are antigenic
CC components of myelin or myelin neurofilaments. Biological samples
CC are assayed for antibodies, especially IGA antibodies, which bind to
CC myelin and/or myelin neurofilaments or an antigenic peptide, such as the
CC present sequence. Any reading in excess of two standard deviations of
CC the healthy controls would indicate a positive response.
XX
SQ Sequence 7 AA;
Query Match 100.0%; Score 36; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 7.8e+05; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;
QY 1 LKRVHEE 7
Db 1 LKRVHEE 7
RESULT 2
ID AAY20612 standard; Protein; 441 AA.
XX AAY20612:
AC
XX
XX 22-JUL-1999 (first entry)
DE Human neurofilament-L wild type protein fragment 2.
XX
XX Human: beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
XX frameshift mutation; age-related disease; neurodegenerative disorder;
XX Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
XX Huntington's disease; multiple sclerosis; alcoholic liver disease;
XX diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
XX ubiquitin B; Apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
XX neurofilament-F; presentin I; presentin II; cellular tumour antigen;
XX glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
XX bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
XX high mobility group protein-C; neuroendocrine specific protein A.
XX
OS Homo sapiens.
XX
XX
XX W09845322-A2.
XX
XX 15-OCT-1998.
XX
XX 02-APR-1998; 98MO-IB00705.
XX
XX 10-APR-1997; 97US-0043163.
XX
XX (UYUT-) RIJNSDUNIV UTRECHT.
XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
XX
XX WPI: 1998-609901/51.
XX N-PSDB: AAX75758.
XX
XX Diagnosing disease by detecting frameshift mutations in RNA or
XX corresponding protein mutations - used to diagnose cancer and
XX neurological diseases, particularly Alzheimer's disease, and also
XX for treatment and prevention with specific ribozymes or wild-type
XX RNA
PS Disclosure; Figure 7; 258pp; English.

XX This invention describes a novel method for the diagnosis of a disease
CC caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
SQ Sequence 441 AA;
Query Match 100.0%; Score 36; DB 19; Length 441;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LKRVHEE 7
Db 262 LKRVHEE 268
RESULT 3
ID AAM93466 standard; Protein; 543 AA.
XX AAM93466:
AC
XX
XX 06-NOV-2001 (first entry)
DE Human polypeptide, SEQ ID NO: 3132.
XX
XX Human: full length cDNA; cDNA synthesis; oligo-capping.
XX
XX Homo sapiens.
XX
XX
XX EP1130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-0114089.
XX
XX 08-JUL-1999; 99JP-0194486.
XX 11-JAN-2000; 2000JP-0118774.
XX 02-MAY-2000; 2000JP-0183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI: 2001-524255/58.
XX N-PSDB: AAK94387.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
XX use in genetic manipulation -
XX
XX Claim 8; SEQ ID NO 3132; 1380pp + sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by

CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.

XX Sequence 543 AA;

Query Match 100.0%; Score 36; DB 22; Length 543;

Best Local Similarity 100.0%; Pred. No. 74;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7

DB 223 LKKVHEE 229

RESULT 4
ABB79475

ID ABB79475 standard; peptide; 15 AA.

XX ABB79475;

XX 23-SEP-2002 (first entry)

XX Human Orc4lp 47.3 N-terminal peptide fragment.

KW Replication start codon initiation recognition compound; Orc4lp;
KW HsORC4L; cancer; cytostatic; HIV infection; anti-HIV; virucide;
KW human; gene therapy.

XX Homo sapiens.

XX CN1331153-A.

PD 16-JAN-2002.

PF 26-JUN-2000; 2000CN-0116778.

PR 26-JUN-2000; 2000CN-0116778.

PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX Mao Y, Xie Y;

DR WPI; 2002-352937/39.

XX Polypeptide-human replication start codon initiation recognition
PT compound subunit Orc4lp (HsORC4L)47.3 and polynucleotide for coding it

XX Example 5; Page 20 (Disclosure); 33pp; Chinese.

CC The present invention relates to novel human replication start
CC codon initiation recognition compound subunit Orc4lp (HsORC4L) 47.3
CC (see ABB79474). This protein and its coding sequence are useful in
CC the treatment of diseases such as cancer and HIV infection. The
CC present sequence is an N-terminal peptide fragment of the protein,
CC which was used in an example from the invention.

XX Sequence 15 AA;

Query Match 91.7%; Score 33; DB 23; Length 15;

Best Local Similarity 85.7%; Pred. No. 7.5;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7

DB 9 LKKVHEE 15

RESULT 5
AAB42280
ID AAB42280 standard; Protein; 319 AA.

XX AAB42280;

XX 08-FEB-2001 (first entry)

DE Human ORFX ORF2044 polypeptide sequence SEQ ID NO:4088.

KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnary; antipapillary; antipapillary; neurotropic; neuroprotective;
KW anticonvulsant; osteopathic; antipapillary; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antineoplastic;
KW antiviral; antibacterial; antifungal; antineoplastic; antihypertensive;
KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive.

XX Homo sapiens.

XX WO200058473-A2.

PD 05-OCT-2000.

PF 31-MAR-2000; 2000WO-US08621.

PR 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0540763.

PA (CURA-) CURAGEN CORP.

XX Shinketsu RA, Leach M;

DR WPI; 2000-602362/57.

XX N-PSDB; AAC76489.

XX Claim 11; Page 3278; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,

XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
XX sequences have activities such as: cytostatic; hepatotropic; vulnary;
XX antipapillary; antipapillary; neurotropic; neuroprotective;
XX osteopathic; anticonvulsant; antipapillary; immunosuppressant;
XX immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
XX antidiabetic; hypotensive; dermatological; immunosuppressive;
XX antineoplastic; antibacterial; antiviral; antineoplastic; antihypertensive;
XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive.

CC AAB42280 standard; Protein; 319 AA.
CC AAB42280;
CC 08-FEB-2001 (first entry)
CC Human ORFX ORF2044 polypeptide sequence SEQ ID NO:4088.
CC Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
CC vulnary; antipapillary; antipapillary; neurotropic; neuroprotective;
CC anticonvulsant; osteopathic; antipapillary; immunosuppressant; cardiant;
CC immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
CC hypotensive; dermatological; immunosuppressive; antineoplastic;
CC antiviral; antibacterial; antifungal; antineoplastic; antihypertensive;
CC antianemic; gene therapy; cancer; proliferative disorder; hypertension;
CC neurodegenerative disorder; osteoarthritis; graft vs host disease;
CC cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
CC cholesterol ester storage; systemic lupus erythematosus; infection;
CC severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
CC allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
CC bone damage; cartilage damage; antiinflammatory disease; coagulation;
CC thrombosis; contraceptive.
CC Claim 11; Page 3278; 5507pp; English.
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
CC antipapillary; antipapillary; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antipapillary; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antineoplastic; antibacterial; antiviral; antineoplastic; antihypertensive;
CC antianemic; gene therapy; cancer; proliferative disorder; hypertension;
CC neurodegenerative disorder; osteoarthritis; graft vs host disease;
CC cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
CC cholesterol ester storage; systemic lupus erythematosus; infection;
CC severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
CC allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
CC bone damage; cartilage damage; antiinflammatory disease; coagulation;
CC thrombosis; contraceptive.

SQ Sequence 319 AA;

Query Match 91.7%; Score 33; DB 21; Length 319;

Best Local Similarity 85.7%; Pred. No. 1.6e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKKVHEE 7

DB 151 LKKLHEE 157

RESULT 6

ABR79474

ID ABR79474 standard; Protein; 430 AA.

AC ABR79474;

DT 23-SEP-2002 (first entry)

DE Human Orc4Lp 47.3 protein.

KW Replication start codon initiation recognition compound; Orc4Lp;

KW HsORC4L; cancer; cytostatic; HIV infection; anti-HIV; virucide;

KW human; gene therapy.

XX Homo sapiens.

XX CN131153-A.

XX 16-JAN-2002.

XX 26-JUN-2000; 2000CN-0116778.

XX 26-JUN-2000; 2000CN-0116778.

XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX Mac Y, Xie Y;

XX WPI: 2002-352937/39.

XX N-PSDB; ABN84122.

XX Polypeptide-human replication start codon initiation recognition

XX compound subunit Orc4Lp (HsORC4L)47.3 and polynucleotide for coding it

XX Claim 1; Page 28-29 (Disclosure); 33pp; Chinese.

XX The present sequence is the protein sequence of novel human

XX replication start codon initiation recognition compound subunit

XX Orc4Lp (HsORC4L) 47.3. This protein and its coding sequence are

XX useful in the treatment of diseases such as cancer and HIV

XX infection.

XX Sequence 430 AA;

QY 1 LKKVHEE 7

DB 9 LKKLHEE 15

RESULT 7

AAW54351

ID AAW54351 standard; protein; 465 AA.

AC AAW54351;

DT 14-AUG-1998 (first entry)

DE Vimentin.

KW Endometrium; hyperplasia; adenocarcinoma; proliferative phase;

KW 2D gel electrophoresis; detection.

XX Homo sapiens.

XX WO9810291-A1.

XX 12-MAR-1998.

XX 05-SEP-1997; 97WO-GB02394.

XX 08-APR-1997; 97GB-0007132.

XX 06-SEP-1996; 96GB-0018600.

XX (CLIN-) CENT CLINICAL & BASIC RES.

XX Byrjalsen I, Fey SJ, Larsen P;

XX WPI: 1998-207057/18.

XX Biochemical markers of human endometrium - useful for, e.g.

XX diagnosis of hyperplasia and adenocarcinoma

XX Disclosure; Page 20; 77pp; English.

XX Proteins AAW54349-W54364 are examples of proteins produced in the

XX endometrium during the hyperplasia, adenocarcinoma or proliferative

XX phase of the endometrium. The presence and quantities of these proteins

XX can be detected using 2D gel electrophoresis comparison of cell lysates.

XX The proteins can be used as biochemical markers to detect the phase of

XX the endometrium and can be measured in body fluids, obviating the need

XX for endometrial biopsies.

SQ Sequence 465 AA;

QY 1 LKKVHEE 7

DB 233 LKKLHEE 239

RESULT 8

ABR77394

ID ABR77394 standard; protein; 465 AA.

AC ABR77394;

DT 11-JUL-2002 (first entry)

DE Human vimentin.

KW Human; dermatological; skin stress; aging; spondin 2; cathepsin L;

KW actin gamma 1; vimentin; fibroblast; skin; cosmetic; pharmaceutical.

XX Homo sapiens.

XX DE10050274-A1.

XX 18-APR-2002.

XX 09-OCT-2000; 2000DE-1050274.

XX 09-OCT-2000; 2000DE-1050274.

XX (HENK) HENKEL KGAA.

XX Petersohn D, Schmitt G, Foerster T;

DR WPI: 2002-373046/41.
 XX In vitro assays for skin stress and skin ageing includes determination
 PT of spondin 2, cathepsin L, actin gamma 1 and vimentin fragments
 PT secreted by skin fibroblasts -
 PS Claim 5: Page 12; 14pp; German.
 CC The invention relates to in vitro methods for the detection of skin
 CC stress and/or skin ageing in humans and animals based on the
 CC determination of spondin 2, cathepsin L, actin gamma 1 or vimentin
 CC fragments secreted by fibroblast from the skin under test. Use of the
 CC methods in a test for potential cosmetics and pharmaceuticals with an
 CC effect on these skin conditions and products containing vimentin
 CC fragments are also included. Products containing vimentin fragments
 CC are effective in the regulation, especially maintenance, of skin
 CC homeostasis.
 SQ Sequence 465 AA;
 XX
 SQ
 Query Match 91.7%; Score 33; DB 23; Length 465;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 LKRVHEE 7
 DB 234 LKRVHEE 240
 RESULT 9
 AAB29635
 ID AAB29635 standard; Protein; 466 AA.
 XX
 AC AAB29635;
 XX
 DT 21-FEB-2001 (first entry)
 DE Human pollinosis-associated gene 795-encoded protein, SEQ ID NO:26.
 XX
 DE Human pollinosis-associated gene 795; vimentin homologue;
 KM Ige; immunoglobulin E; cedar pollen allergy; T-cell; reduced expression;
 KM detection; diagnosis; drug screening; allergic disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200065050-A1.
 XX
 PD 02-NOV-2000.
 XX
 PD 26-APR-2000; 2000WO-JP02734.
 XX
 PF 27-APR-1999; 99JP-0120494.
 XX
 PR (GENO-) GENOX RES INC.
 PA (EISA) EISAI CO LTD.
 XX
 PI Nagasu T, Sugita Y, Kashiwabara T, Oshida T, Odayashi M, Gunji S;
 PI Odayashi I, Imai Y, Yoshida N, Ogawa K, Matsui K, Takahashi E;
 PI Yokoi A;
 DR WPI: 2000-687343/67.
 DR N-PSDB; AAC64226.
 XX
 PT Pollinosis-associated gene 795 undergoing significantly low expression
 PT in subjects with high cedar pollen-specific IgE levels; useful in
 PT diagnosis of allergic diseases and screening drug candidates -
 PS Page 64-67; Claim 13; 73pp; Japanese.
 CC The invention relates to the human pollinosis-associated gene 795 which
 CC exhibits significantly reduced expression in the T-cells of individuals
 CC with high cedar pollen-specific IgE (immunoglobulin E) levels. The gene
 CC was isolated from T-cells from individuals allergic to cedar pollen using

CC the differential display method. Pollinosis-associated gene 795 has
 CC homology with the human vimentin gene. The invention also relates also
 CC relates to the protein encoded by pollinosis gene 795; to expression
 CC constructs and host cells comprising pollinosis-associated gene 795
 CC nucleic acids; pollinosis-associated gene 795 primers and probes;
 CC antibodies against the protein encoded by the gene; methods of detection
 CC of pollinosis-associated gene 795 nucleic acids; and a method of
 CC diagnosis of allergic diseases via the detection of pollinosis-associated
 CC gene 795 nucleic acids. The invention additionally encompasses methods of
 CC screening drug candidates for the treatment of allergic disease by
 CC measuring the expression of pollinosis-associated gene 795 in pollen
 CC antigen-stimulated T-cells in the presence of a test compound relative to
 CC a control. Pollinosis-associated gene 795 is useful in the diagnosis of
 CC allergic diseases and in the screening of drug candidates for the
 CC treatment of such diseases. The present sequence represents a
 CC protein encoded by human pollinosis-associated gene 795.
 SQ Sequence 466 AA;
 XX
 SQ
 Query Match 91.7%; Score 33; DB 21; Length 466;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 LKRVHEE 7
 DB 234 LKRVHEE 240
 RESULT 10
 AAY92335
 ID AAY92335 standard; Protein; 466 AA.
 XX
 AC AAY92335;
 XX
 DT 21-AUG-2000 (first entry)
 DE Human vimentin.
 XX
 DE Human vimentin.
 KM NIK1 Interacting Protein; vimentin; protein complex; cytoskeletal;
 KM antiviral; neuroprotective; cardiant.
 XX
 OS Homo sapiens.
 XX
 PN WO200020448-A2.
 XX
 PD 13-APR-2000.
 XX
 PD 06-OCT-1999; 99WO-US23314.
 XX
 PF 06-OCT-1998; 98US-0167206.
 XX
 PR (CURA-) CURAGEN CORP.
 PA
 XX
 PI Nandabalan K, Schulz VP, Yang M;
 DR WPI: 2000-303742/26.
 DR N-PSDB; AAA09307.
 XX
 PT New complex of a NIK1 protein and a NIK1 protein-interacting protein,
 PT useful for treating cancer, hyperproliferative disorder,
 PT neurodegenerative disorder, cardiomyopathies, viral infections and
 PT metabolic disorders
 XX
 PS Example 1; Page 145-147; 172pp; English.
 CC AAY92331-37 were isolated in a modified yeast two hybrid system using
 CC NIK1 protein as "bait". These are known sequences which are NIK1
 CC interacting proteins. The invention concerns purified complexes of a
 CC NIK1 protein and a NIK1 protein-interacting protein, where the
 CC interacting protein is chosen from TrkA, protein phosphatase 1alpha,
 CC 14-3-3epsilon, alpha-tropomyosin, vimentin, p0071, Inl-1, IP-1, IP-2,
 CC IP-3, IP-4, or IP-5. NIK1 (also referred to as Nek2) is a human
 CC homologue of the Aspergillus nidulans mitotic regulator, NIMA kinase.

CC NIK1 is a serine/threonine-specific kinase and is thought to play a key
 CC role in cell-cycle events leading to the onset of mitosis. The
 CC complexes, their derivatives and NIK1 or NIK1-IP protein and DNA
 CC sequences, etc. are useful for treating or preventing a disease or
 CC disorder involving aberrant levels of the complex or protein. Such
 CC disorders include cancer, hyperproliferative disorders, such
 CC neurodegenerative disorders, cardiomyopathies, viral infections and
 CC metabolic disorders.
 CC
 SQ Sequence 466 AA;

Query Match 91.7%; Score 33; DB 21; Length 466;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
 |||:||||
 Db 234 LKRLHEE 240

RESULT 11

AAB6348
 ID AAB6348 standard; peptide; 466 AA.

AC AAB6348;

DT 05-APR-2001 (first entry)

DE Human vimentin.

XX Vimentin; caspase cleavage site; apoptosis; antibody; Basedow's disease;
 KM systemic lupus erythematosus; autoimmune hemolytic anaemia; AIDS; human.

OS Homo sapiens.

XX EP1067142-A1.

XX 10-JAN-2001.

XX 07-JUL-2000; 2000EP-0305736.

XX 07-JUL-1999; 99JP-0193235.

XX (RIKE) RIKEN KK.

PI Morishima N, Nakanishi K, Shibata T;

DR WPI; 2001-149349/16.

XX New antibody reacting with a cleavage product of vimentin but not with
 PT the intact vimentin, useful for detecting apoptosis and the quantity of
 PT cleavage product of vimentin or as a reagent for immunohistochemical
 PT staining
 CC Disclosure: Page 10-12; 23pp; English.

CC The present invention provides an antibody which reacts with a cleavage
 CC product of vimentin but not with the intact protein. This can be used to
 CC detect apoptosis, which may then be used as an indicator of the
 CC progression of diseases such as systemic lupus erythematosus, autoimmune
 CC haemolytic anaemia, Basedow's disease and acquired immunodeficiency
 CC syndrome (AIDS).
 CC
 SQ Sequence 466 AA;

Query Match 91.7%; Score 33; DB 22; Length 466;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
 |||:||||
 Db 234 LKRLHEE 240

RESULT 12

AAU87694

ID AAU87694 standard; Protein; 466 AA.

XX AAU87694;

DT 21-MAY-2002 (first entry)

DE Human pancreatic tumour protein #6.

XX Human; pancreatic tumour protein; immune response; pancreatic cancer;
 KM development of cancer; cancer progression; cytostatic.
 XX
 OS Homo sapiens.

PN WO200212331-A2.

PD 14-FEB-2002.

PF 06-AUG-2001; 2001WO-US24619.

PR 07-AUG-2000; 2000US-223130P.

PR 30-JAN-2001; 2001US-265447P.

PR 15-MAY-2001; 2001US-291201P.

PA (CORI-) CORIXA CORP.

PI Pyle RA, Xu J, Kalos MD;

DR WPI; 2002-241741/29.

DR N-PSDB; ABR44209.

XX Novel polynucleotide encoding pancreatic tumour polypeptides, useful in
 PT pharmaceutical compositions, e.g. vaccines, for treating pancreatic
 PT cancers
 CC Claim 2; Page 165-167; 167pp; English.

XX The present invention relates to the isolation of cDNA sequences
 CC (ABR44061-ABR44209) encoding human pancreatic tumour proteins. The
 CC polynucleotide sequences encoding human pancreatic tumour proteins are
 CC useful for stimulating an immune response in a patient and treating
 CC pancreatic cancer in a patient. A host cell that expresses these
 CC polynucleotides is useful for determining the presence of cancer in a
 CC patient. A composition comprising the polynucleotide, its encoded
 CC protein, or an antibody that binds to the protein may be used in the
 CC diagnosis, prevention and/or treatment of diseases, particularly in
 CC pancreatic cancer. The sequences of the invention are also useful in
 CC pharmaceutical compositions, e.g. vaccines, for the diagnosis and
 CC treatment of pancreatic cancer. Such compositions may be useful for
 CC inhibiting the development of cancer in a patient, or as markers for
 CC the progression of cancer. The polynucleotide sequences may also be used
 CC as probes or primers for nucleic acid hybridisation assays.
 CC AAU87689-AAU87694 represent human pancreatic tumour proteins.
 CC
 SQ Sequence 466 AA;

Query Match 91.7%; Score 33; DB 23; Length 466;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
 |||:||||
 Db 234 LKRLHEE 240

RESULT 13

AAB95511

ID AAB95511 standard; Protein; 475 AA.

XX AAB95511;

DT 26-JUN-2001 (first entry)
 XX Human protein sequence SEQ ID NO:18077.
 DE Human protein sequence SEQ ID NO:18077.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 XX EP1074617-A2.
 XX 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 XX 27-AUG-1999; 99JP-0300253.
 XX 11-JAN-2000; 2000JP-0118776.
 XX 02-MAY-2000; 2000JP-0183767.
 XX 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 18077; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 475 AA;
 XX
 Query Match 91.7%; Score 33; DB 22; Length 475;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX 1 LKAYHEE 7
 XX |||:||||
 Db 100 LKRLHEE 106
 XX
 RESULT 14
 AAB94625
 ID AAB94625 standard; Protein; 541 AA.
 XX
 AC AAB94625;

XX 26-JUN-2001 (first entry)
 XX Human protein sequence SEQ ID NO:15483.
 DE Human protein sequence SEQ ID NO:15483.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 XX EP1074617-A2.
 XX 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 XX 27-AUG-1999; 99JP-0300253.
 XX 11-JAN-2000; 2000JP-0118776.
 XX 02-MAY-2000; 2000JP-0183767.
 XX 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 15483; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 541 AA;
 XX
 Query Match 91.7%; Score 33; DB 22; Length 541;
 Best Local Similarity 85.7%; Pred. No. 2.7e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX 1 LKAYHEE 7
 XX |||:||||
 Db 100 LKRLHEE 106
 XX
 RESULT 15
 AAM25417
 ID AAM25417 standard; Protein; 545 AA.
 XX

AC AAM25417;
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:932.
 XX
 KM Human: cancer; HIV infection; human immunodeficiency virus;
 KM antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KM antibacterial; endocrine; cardiant; central nervous system; virucide;
 KM anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
 KM antiaggregant; haemostatic; vulnery; antilucer; osteopathic; eczema;
 KM dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
 KM neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KM antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KM genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KM thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
 KM allergic rhinitis; diabetes; multiple sclerosis; depression;
 KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KM neurological disorder.
 XX
 OS Homo sapiens.
 XX
 OS WO200153455-A2.
 XX
 PN 26-JUL-2001.
 XX
 PD 22-DEC-2000; 2000WO-US35017.
 XX
 PF 23-DEC-1999; 99US-0471275.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX
 PI Tang YF, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-457603/49.
 DR N-PSDB: AAM99358.
 XX
 XX Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX
 XX Claim 20; Page 201; 1217pp; English.
 PS
 XX AAM99166 to AAM99904 encode the human proteins given in AAM25225 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
 CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
 CC antilucer; osteopathic; dermatological; antiallergic; antiasthmatic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 CC
 XX Sequence 545 AA;

Query Match 91.7%; Score 33; DB 22; Length 545;
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
 |||:||||
 Db 150 LKRLHEE 156

Search completed: November 13, 2002, 13:22:06
 Job time : 36.7021 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 10.1277 seconds
(without alignments)
20.336 Million cell updates/sec

Title: US-09-856-086-2

Perfect score: 36

Sequence: 1 LKKVHE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	160	2	US-08-726-306A-183
2	33	91.7	466	4	US-09-610-401-3
3	33	91.7	3135	1	US-08-323-170B-2
4	33	91.7	3135	4	US-08-954-441-2
5	30	83.3	420	2	US-08-592-126-142
6	30	83.3	466	4	US-09-610-401-4
7	30	83.3	746	4	US-09-149-934-4
8	30	83.3	804	4	US-09-134-001C-5218
9	29	80.6	106	2	US-08-893-042-1
10	29	80.6	255	2	US-08-685-992-16
11	29	80.6	255	2	US-09-144-925-16
12	29	80.6	400	1	US-07-730-953-2
13	29	80.6	432	2	US-08-705-660-18
14	29	80.6	432	3	US-08-989-045-18
15	29	80.6	3079	5	PCT-US94-00198-4
16	28	77.8	68	4	US-09-911-882-33
17	28	77.8	68	4	US-09-911-882-33
18	28	77.8	69	4	US-08-965-762-33
19	28	77.8	147	1	US-08-464-342-2
20	28	77.8	147	1	US-08-305-520-2
21	28	77.8	147	2	US-08-464-604A-2
22	28	77.8	147	2	US-08-875-272-2
23	28	77.8	147	2	US-08-486-663A-2
24	28	77.8	147	2	US-08-903-396-2
25	28	77.8	147	2	US-08-247-904B-2
26	28	77.8	147	3	US-08-895-601-8
27	28	77.8	147	3	US-08-767-942A-2

28	28	77.8	154	4	US-09-134-001C-4529	Sequence 4529, Ap
29	28	77.8	280	4	US-09-264-419C-2	Sequence 2, Appl
30	28	77.8	289	4	US-09-071-035-28	Sequence 28, Appl
31	28	77.8	308	4	US-09-071-035-26	Sequence 26, Appl
32	28	77.8	390	4	US-09-419-459-4	Sequence 4, Appl
33	28	77.8	394	4	US-09-419-459-10	Sequence 10, Appl
34	28	77.8	355	2	US-08-007-107-2	Sequence 2, Appl
35	28	77.8	582	4	US-09-419-459-2	Sequence 2, Appl
36	28	77.8	619	4	US-09-134-001C-4248	Sequence 4248, Ap
37	28	77.8	984	4	US-08-764-870-15	Sequence 15, Appl
38	28	77.8	984	4	US-08-980-115-15	Sequence 15, Appl
39	27	75.0	7	4	US-09-172-045-4	Sequence 4, Appl
40	27	75.0	56	2	US-08-323-449B-3	Sequence 3, Appl
41	27	75.0	56	2	US-08-323-449B-4	Sequence 4, Appl
42	27	75.0	56	2	US-08-485-981-3	Sequence 3, Appl
43	27	75.0	56	2	US-08-485-981-4	Sequence 4, Appl
44	27	75.0	56	2	US-08-867-087B-4	Sequence 4, Appl
45	27	75.0	56	2	US-08-867-087B-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-726-306A-183
Sequence 183, Application US/08726306A
Patent No. 5958684
GENERAL INFORMATION:
APPLICANT: van Leeuwen, Frederik Willem
APPLICANT: Burbach, Johannes Peter Henri
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1 Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,306A
FILING DATE: 02-Oct-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080, 4
FILING DATE: 02-Oct-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/009,832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 183:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-726-306A-183
Query Match 100.0%; Score 36; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
DB 9 LKKVHEE 15

RESULT 2
US-09-610-401-3

Sequence 3, Application US/09610401

Patent No. 6417336

GENERAL INFORMATION:

APPLICANT: MORISHIMA, No. 6417336unh1ro,

APPLICANT: NAKASHIMA, Keiko,

APPLICANT: SHIBATA, Takehiko

TITLE OF INVENTION: Antibody against cleavage product of vimentin

FILE REFERENCE: 522.1004

CURRENT APPLICATION NUMBER: US/09/610,401

PRIOR FILING DATE: 2000-07-05

PRIOR APPLICATION NUMBER: JP 11-193235

NUMBER OF SEQ ID NOS: 4

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 3

LENGTH: 466

TYPE: PRT

ORGANISM: Homo sapiens

US-09-610-401-3

Query Match 91.7%; Score 33; DB 4; Length 466;
Best Local Similarity 85.7%; Pred. No. 69;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
DB 234 LKKIHEE 240

RESULT 3
US-08-323-170B-2

Sequence 2, Application US/08323170B

Patent No. 5733772

GENERAL INFORMATION:

APPLICANT: Williamson, Kim C.

APPLICANT: Kaslow, David C.

TITLE OF INVENTION: Cloning and Expression of Plasmidium

TITLE OF INVENTION: falciptarum Transmission-Blocking Target Antigen, pfs230

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/323,170B

CLASSIFICATION: 424

FILING DATE: 13-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/010,409

FILING DATE: 29-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Quine, Jonathan A.

REGISTRATION NUMBER: P-41,261

REFERENCE/DOCKET NUMBER: 015280-113100US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

;; INFORMATION FOR SEQ ID NO: 2;
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3135 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-323-170B-2

Query Match 91.7%; Score 33; DB 1; Length 3135;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
DB 2555 IKKIHEE 2561

RESULT 4
US-08-954-441-2

Sequence 2, Application US/08954441

Patent No. 6316000

GENERAL INFORMATION:

APPLICANT: Williamson, Kim C.

APPLICANT: Kaslow, David C.

TITLE OF INVENTION: Cloning and Expression of Plasmidium

TITLE OF INVENTION: falciptarum Transmission-Blocking Target Antigen, pfs230

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/954,441

FILING DATE: 20-OCT-1997

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/323,170

FILING DATE: 13-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/010,409

FILING DATE: 29-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Einhorn, Gregory P.

REGISTRATION NUMBER: 38,440

REFERENCE/DOCKET NUMBER: 015280-113110US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 3135 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-954-441-2

Query Match 91.7%; Score 33; DB 4; Length 3135;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
DB 2555 IKKIHEE 2561

RESULT 5
US-08-592-126-142
; Sequence 142, Application US/08592126
; Patent No. 5821091
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Transcripts Encoding Immunomodulatory
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 151
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592.126
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0860
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 142:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 420 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: CDC.pcp
; US-08-592-126-142

Query Match 83.3%; Score 30; DB 2; Length 420;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
|||:|
DB 373 LKRVHGE 379

RESULT 6
US-09-610-401-4
; Sequence 4, Application US/09610401
; Patent No. 6417336
; GENERAL INFORMATION:
; APPLICANT: MORISHIMA, No. 6417336uhito,
; APPLICANT: NAKANISHI, Keiko,
; APPLICANT: SHIBATA, Takehiko
; TITLE OF INVENTION: Antibody against cleavage product of vimentin
; FILE REFERENCE: 522.1004
; CURRENT APPLICATION NUMBER: US/09/610.401
; CURRENT FILING DATE: 2000-07-05
; PRIOR APPLICATION NUMBER: JP 11-193235
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 466
; TYPE: PRT

; ORGANISM: Mus sp.
US-09-610-401-4
Query Match 83.3%; Score 30; DB 4; Length 466;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
|||:|
DB 234 LKRIHDE 240

RESULT 7
US-09-149-934-4
; Sequence 4, Application US/09149934B
; Patent No. 6139837
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: ATP-DEPENDENT RNA HELICASE PROTEIN
; FILE REFERENCE: PF-0338-1 DIV
; CURRENT APPLICATION NUMBER: US/09/149.934B
; CURRENT FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 08/892,256
; EARLIER FILING DATE: 1997-07-11
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 746
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: GenBank ID No. 6139837 1707046
; US-09-149-934-4

Query Match 83.3%; Score 30; DB 4; Length 746;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
|||:|
DB 340 LKRIHGE 346

RESULT 8
US-09-134-001C-5218
; Sequence 5218, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5218
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-5218

Query Match 83.3%; Score 30; DB 4; Length 804;
Best Local Similarity 85.7%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
||| |||
Db 440 LKREHEE 446

RESULT 9

US-08-893-042-1
Sequence 1, Application US/08893042
Patent No. 5906923
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL ATPASE INHIBITOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893,042
FILING DATE: Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0134 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-4166
TELEFAX: 415-855-0555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-08-893-042-1

Query Match 80.6%; Score 29; DB 2; Length 106;
Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
||| |||
Db 70 LKRVHEE 76

RESULT 10

US-08-685-992-16
Sequence 16, Application US/08685992
Patent No. 5912138
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas
TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Millitia Drive
CITY: Lexington

STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,992
FILING DATE: 25-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: GSHL96-03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
TELEX:

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-685-992-16

Query Match 80.6%; Score 29; DB 2; Length 255;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
||| |||
Db 31 LKRVHEE 37

RESULT 11

US-09-144-925-16
Sequence 16, Application US/09144925
Patent No. 5951979
GENERAL INFORMATION:
APPLICANT: Tonks, Nicholas
TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02421-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,925
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/685,992
FILING DATE: July 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: CSHL96-032
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-6240
TELEFAX: 781-861-9540
TELEX:
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-144-925-16

Query Match 80.6%; Score 29; DB 2; length 255;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
||| |||
DB 31 LKPIHHE 37

RESULT 12
US-07-730-953-2
Sequence 2, Application US/07730953
Patent No. 5288614
GENERAL INFORMATION:
APPLICANT: BODENMULLER, Heinz
APPLICANT: DESSAUER, Andreas
TITLE OF INVENTION: METHOD FOR THE DETECTION OF MALIGNANT
TITLE OF INVENTION: DISEASES
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKALDO, Marmelstein, Murray & Oram
STREET: 655 Fifteenth Street N.W. Suite 330
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/730,953
CLASSIFICATION: 435
FILING DATE: 19910723
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE P 39 42 999.7
FILING DATE: 21-DEC-1989
ATTORNEY/AGENT INFORMATION:
NAME: KILTS, Monica C.
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: P564-1119
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)638-5000
TELEFAX: (202)638-4810
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-730-953-2

Query Match 80.6%; Score 29; DB 1; length 400;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
||| |||

DB 214 LKKNHEE 220

RESULT 13
US-08-705-660-18
Sequence 18, Application US/08705660
Patent No. 5858683
GENERAL INFORMATION:
APPLICANT: KEESEE, SUSAN
APPLICANT: OBAR, ROBERT
APPLICANT: WU, YING-JYE
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/705,660
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: GREENHALGH, DUNCAN A
REGISTRATION NUMBER: 38,678
REFERENCE/DOCKET NUMBER: MTP-023 (8395/27)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-705-660-18

Query Match 80.6%; Score 29; DB 2; length 432;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
||| |||
DB 218 LKKNHEE 224

RESULT 14
US-08-989-045-18
Sequence 18, Application US/08989045
Patent No. 6027905
GENERAL INFORMATION:
APPLICANT: KEESEE, SUSAN
APPLICANT: OBAR, ROBERT
APPLICANT: WU, YING-JYE
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
TITLE OF INVENTION: DETECTION OF CERVICAL CANCER
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault
STREET: 125 High St.
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/989,045
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: GREENHALGH, DUNCAN A
REGISTRATION NUMBER: 38,678
REFERENCE/DOCKET NUMBER: MTP-023 (8395/27).
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-989-045-18

Query Match 80.6%; Score 29; DB 3; Length 432;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 LKKVHEE 7
DB 218 LKKNHEE 224

RESULT 15
PCT-US94-00198-4
Sequence 4, Application PC/TUS9400198
GENERAL INFORMATION:
APPLICANT: Schering Corp.
TITLE OF INVENTION: RAS Associated GAP Proteins
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 1 Giraldo Farms
CITY: Madison
STATE: New Jersey
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 6.0.8
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/00198
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/004,824
FILING DATE: 15-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G.
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: DX0352 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)822-7255
TELEFAX: (201)822-7039
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3079 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
PCT-US94-00198-4

Query Match 80.6%; Score 29; DB 5; Length 3079;
Best Local Similarity 71.4%; Pred. No. 1.9e+03;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 LKKVHEE 7
DB 3019 LKKNHEE 3025

Search completed: November 13, 2002, 13:18:05
Job time : 11.1277 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2002, 13:18:18 ; Search time 4.76596 Seconds
(without alignments)
22.121 Million cell updates/sec

Title: US-09-856-086-2
Perfect score: 36
Sequence: 1 LKRVHEE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PC1_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	91.7	466	10 US-09-923-779-155	Sequence 155, App
2	33	91.7	466	12 US-10-152-647-3	Sequence 3, Appl
3	32	88.9	400	10 US-09-881-752A-206	Sequence 206, App
4	30	83.3	64	10 US-09-864-761-39889	Sequence 39889, A
5	30	83.3	336	10 US-09-764-898-281	Sequence 281, App
6	30	83.3	466	12 US-10-152-647-4	Sequence 4, Appl
7	29	80.6	39	10 US-09-864-761-42664	Sequence 42664, A
8	29	80.6	81	10 US-09-935-145-5	Sequence 5, Appl
9	29	80.6	106	10 US-09-273-135-1	Sequence 1, Appl
10	29	80.6	106	10 US-09-974-216-1	Sequence 1, Appl
11	29	80.6	123	10 US-09-925-301-1456	Sequence 1456, Ap
12	29	80.6	223	10 US-09-816-494-5	Sequence 5, Appl
13	29	80.6	313	10 US-09-788-626-17	Sequence 17, Appl
14	29	80.6	400	10 US-09-922-217-1115	Sequence 115, Ap
15	29	80.6	432	10 US-09-919-172-9	Sequence 9, Appl
16	29	80.6	456	10 US-09-919-172-31	Sequence 31, Appl
17	29	80.6	549	10 US-09-764-864-1131	Sequence 1131, Ap
18	29	80.6	618	10 US-09-925-300-1381	Sequence 181, Ap
19	28	77.8	68	10 US-09-911-888-33	Sequence 33, Appl

20	28	77.8	134	10 US-09-864-761-34684	Sequence 34684, A
21	28	77.8	256	10 US-09-925-299-992	Sequence 992, App
22	28	77.8	280	9 US-09-905-291A-325	Sequence 325, App
23	28	77.8	280	10 US-09-909-320-325	Sequence 325, App
24	28	77.8	280	10 US-09-909-088B-325	Sequence 325, App
25	28	77.8	280	12 US-10-052-586-30	Sequence 30, Appl
26	28	77.8	398	10 US-09-729-674-146	Sequence 146, Appl
27	28	77.8	422	10 US-09-779-307-20	Sequence 20, Appl
28	28	77.8	428	10 US-09-779-307-21	Sequence 21, Appl
29	28	77.8	430	10 US-09-779-307-19	Sequence 19, Appl
30	28	77.8	435	10 US-09-866-582-33	Sequence 33, Appl
31	28	77.8	489	10 US-09-815-242-5850	Sequence 5850, Ap
32	28	77.8	581	12 US-10-074-527-2	Sequence 2, Appl
33	28	77.8	600	10 US-09-975-901-2	Sequence 2, Appl
34	28	77.8	601	10 US-09-815-242-5638	Sequence 5638, Ap
35	28	77.8	607	10 US-09-815-242-13368	Sequence 12368, A
36	28	77.8	607	10 US-09-815-242-13368	Sequence 12368, A
37	28	77.8	611	10 US-09-815-242-12111	Sequence 12111, A
38	28	77.8	981	10 US-09-815-242-12111	Sequence 12111, A
39	27	75.0	10	12 US-10-026-001-7	Sequence 7, Appl
40	27	75.0	35	10 US-09-864-761-42517	Sequence 42517, A
41	27	75.0	62	10 US-09-864-761-47501	Sequence 47501, A
42	27	75.0	63	10 US-09-864-761-35183	Sequence 35183, A
43	27	75.0	117	10 US-09-764-847-907	Sequence 907, App
44	27	75.0	135	10 US-09-925-301-1529	Sequence 1529, Ap
45	27	75.0	200	10 US-09-785-738A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-923-779-155
Sequence 155, Application US/09923779
Patent No. US20020076721A1
GENERAL INFORMATION:
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiaochun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.553
CURRENT APPLICATION NUMBER: US/09/923, 779
CURRENT FILING DATE: 2001-08-06
NUMBER OF SEQ ID NOS: 155
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 155
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-09-923-779-155

Query Match 91.7%; Score 33; DB 10; Length 466;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 234 LKRVHEE 240

RESULT 2
US-10-152-647-3
Sequence 3, Application US/10152647
Patent No. US20020137110A1
GENERAL INFORMATION:
APPLICANT: MORISHIMA, No. US20020137110A1unlfr.
APPLICANT: NAKAMISHI, Keiko.
APPLICANT: SHIBATA, Takehiko
TITLE OF INVENTION: Antibody against cleavage product of vimentin
FILE REFERENCE: 522.1004
CURRENT APPLICATION NUMBER: US/10/152,647
CURRENT FILING DATE: 2002-05-21

PRIOR APPLICATION NUMBER: US/09/610,401
PRIOR FILING DATE: 2000-07-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-10-152-647-3

Query Match 91.7%; Score 33; DB 12; Length 466;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
|||:|
Db 234 LKRLHEE 240

RESULT 3
US-09-881-752A-206
Sequence 206, Application US/09881752A
Patent No. US20020115078A1
GENERAL INFORMATION:
APPLICANT: Kleantous, Harold
APPLICANT: Al-Garawi, Amal
APPLICANT: Miller, Charles
APPLICANT: Tomb, Jean-Francois
APPLICANT: Oomen, Raymond P.
TITLE OF INVENTION: Identification of Polynucleotides
TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
FILE OF INVENTION: Genome
FILE REFERENCE: 06133/041002
CURRENT APPLICATION NUMBER: US/09/881,752A
CURRENT FILING DATE: 2001-06-15
PRIOR APPLICATION NUMBER: US 08/833,457
PRIOR FILING DATE: 1997-04-01
NUMBER OF SEQ ID NOS: 370
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 206
LENGTH: 400
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-881-752A-206

Query Match 88.9%; Score 32; DB 10; Length 400;
Best Local Similarity 85.7%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
|||:|
Db 34 LKRTHEE 40

RESULT 4
US-09-864-761-39889
Sequence 39889, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecolica-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
SEQ ID NO 39889
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004775.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 3.6
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 3.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 6.9
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 3.7
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 3.2
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 4
OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL - 3.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 3.5
OTHER INFORMATION: SWISSPROT HIT: Q92599, EVALUATE 2.00e-19
OTHER INFORMATION: EST_HUMAN HIT: BE748158.1, EVALUATE 2.00e-18
US-09-864-761-39889

Query Match 83.3%; Score 30; DB 10; Length 64;
Best Local Similarity 71.4%; Pred. No. 9.4;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
|||:|
Db 8 LKRVHEE 14

RESULT 5
US-09-764-898-281
Sequence 281, Application US/09764898
Patent No. US20020090673A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PJ201
CURRENT APPLICATION NUMBER: US/09/764,898
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 311
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 281
LENGTH: 336
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (126)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (168)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (186)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (299)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (318)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (329)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-898-281

Query Match
Best Local Similarity 83.3%; Score 30; DB 10; Length 336;
Best Local Similarity 71.4%; Pred. No. 53;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 269 LKRVHDE 275

RESULT 6
US-10-152-647-4
Sequence 4, Application US/10152647
Patent No. US20020137110A1
GENERAL INFORMATION:
APPLICANT: MORISHIMA, No. US20020137110A1uhhro,
APPLICANT: NAKAMISHI, Keiichi,
APPLICANT: SHIBATA, Takehiko,
TITLE OF INVENTION: Antibody against cleavage product of vimentin
FILE REFERENCE: 522.1004
CURRENT APPLICATION NUMBER: US/10/152,647
CURRENT FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US/09/610,401
PRIOR FILING DATE: 2000-07-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 466
TYPE: PRT
ORGANISM: Mus sp.
US-10-152-647-4

Query Match
Best Local Similarity 83.3%; Score 30; DB 12; Length 466;
Best Local Similarity 71.4%; Pred. No. 74;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 234 LKRVHDE 240

RESULT 7
US-09-864-761-42664
Sequence 42664, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmlca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine, vers. 1.1
SEQ ID NO 42664
LENGTH: 39
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000246.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.61
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.67
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.7
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.86
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.65
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.7
OTHER INFORMATION: EST_HUMAN HIT: AA233606.1, EVALUATE 3.00e-15
US-09-864-761-42664

Query Match
Best Local Similarity 80.6%; Score 29; DB 10; Length 39;
Best Local Similarity 57.1%; Pred. No. 8.8;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 15 MOKIHEE 21

RESULT 8
US-09-935-145-5

Sequence 5, Application US/09935145
Patent No. US2002091081A1
GENERAL INFORMATION:
APPLICANT: Papathanassiou, Adonias E.
TITLE OF INVENTION: Compositions and methods for inhibiting angiogenesis
FILE REFERENCE: Sequence Listing-09/935,145
CURRENT APPLICATION NUMBER: US/09/935,145
CURRENT FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/227,152
PRIOR FILING DATE: 2000-08-22
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 81
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-935-145-5

Query Match 80.6%; Score 29; DB 10; Length 81;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
||| |||
DB 45 LKHHHEE 51

RESULT 9
US-09-273-135-1
Sequence 1, Application US/09273135
Patent No. US20020034510A1
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL ATPASE INHIBITOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/273,135
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/725,025
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0134 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: Consensus
US-09-273-135-1

Query Match 80.6%; Score 29; DB 10; Length 106;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
||| |||
DB 70 LKHHHEE 76

RESULT 10
US-09-974-216-1
Sequence 1, Application US/09974216
Patent No. US20020098561A1
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL ATPASE INHIBITOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/974,216
FILING DATE: 09-Oct-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/273,135
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0134 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 106 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: <Unknown>
CLONE: Consensus
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-974-216-1

Query Match 80.6%; Score 29; DB 10; Length 106;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
||| |||
DB 70 LKHHHEE 76

RESULT 11
US-09-925-301-1456
Sequence 1456, Application US/09925301
Patent No. US20020052308A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA106
CURRENT APPLICATION NUMBER: US/09/925,301
CURRENT FILING DATE: 2001-08-10

```

; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 1456
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (131)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (137)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1456
```

```

Query Match      80.6%; Score 29; DB 10; Length 143;
Best Local Similarity 85.7%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 LKKVHEE 7
        ||| |||
Db      90 LKKHHEE 96
```

```

RESULT 12
US-09-816-494-5
; Sequence 5; Application US/09816494
; Patent No. US20020034807A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel A.
; TITLE OF INVENTION: 38692 AND 21117 NOVEL DUAL SPECIFICITY
; FILE REFERENCE: 10448-030002
; CURRENT APPLICATION NUMBER: US/09/816,494
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,858
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-494-5
```

```

Query Match      80.6%; Score 29; DB 10; Length 223;
Best Local Similarity 71.4%; Pred. No. 54;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 LKKVHEE 7
        ||| |||
Db      202 LKRTHEE 208
```

```

RESULT 13
US-09-788-626-17
; Sequence 17; Application US/09788626
; Patent No. US2002009762A1
; GENERAL INFORMATION:
; APPLICANT: Flint, Deborah J.
; APPLICANT: Cool, Deborah E.
; TITLE OF INVENTION: IMPROVED ASSAY FOR PROTEIN TYROSINE
; FILE REFERENCE: 200125,401
; CURRENT APPLICATION NUMBER: US/09/788,626
; CURRENT FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
```

```

; LENGTH: 313
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-788-626-17
```

```

Query Match      80.6%; Score 29; DB 10; Length 313;
Best Local Similarity 71.4%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 LKKVHEE 7
        ||| |||
Db      31 LKPIHEE 37
```

```

RESULT 14
US-09-922-217-1115
; Sequence 1115; Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Iodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yuguang
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121,471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1115
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1115
```

```

Query Match      80.6%; Score 29; DB 10; Length 400;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 LKKVHEE 7
        ||| |||
Db      214 LKKHHEE 220
```

```

RESULT 15
US-09-919-172-9
; Sequence 9; Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Farris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 1454852CD1
US-09-919-172-9

Query Match 80.6%; Score 29; DB 10; Length 432;
Best Local Similarity 85.7%; Pred No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
||| |||
Db 218 LKKNHHE 224

Search completed: November 13, 2002, 13:40:17
Job time : 4.90881 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:16:18 ; Search time 140.447 Seconds
(without alignments)
32.134 Million cell updates/sec

Title: US-09-856-086-2

Perfect score: 36

Sequence: 1 LKRVHEE 7

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCRTUS.COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06.COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07.COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081.COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082.COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083.COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084.COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085.COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086.COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087.COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088.COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089.COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090.COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091.COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092.COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093.COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094.COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095.COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096.COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097.COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098.COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099.COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100.COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101.COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102.COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	448	21	US-09-791-537-94919
2	36	100.0	450	25	US-10-108-2604-4090
3	36	100.0	455	21	US-09-791-537-32570
4	36	100.0	451	21	US-09-791-537-94920
5	36	100.0	462	21	US-09-791-537-94921
6	36	100.0	463	21	US-09-791-537-73564

Result No.	Score	Query Match	Length	ID	Description
7	36	100.0	469	19	US-09-538-092-948
8	36	100.0	469	21	US-09-791-537-62363
9	36	100.0	469	21	US-09-791-537-73566
10	36	100.0	469	21	US-09-791-537-73566
11	36	100.0	469	21	US-09-791-537-132613
12	36	100.0	469	21	US-09-791-537-132613
13	36	100.0	469	21	US-09-791-537-132613
14	36	100.0	470	21	US-09-791-537-46962
15	36	100.0	470	21	US-09-791-537-62106
16	36	100.0	470	21	US-09-791-537-65178
17	36	100.0	470	21	US-09-791-537-86979
18	36	100.0	470	21	US-09-791-537-86980
19	36	100.0	470	21	US-09-791-537-86983
20	36	100.0	470	21	US-09-791-537-86983
21	36	100.0	471	21	US-09-791-537-1279
22	36	100.0	471	21	US-09-791-537-1279
23	36	100.0	471	21	US-09-791-537-46964
24	36	100.0	471	21	US-09-791-537-92908
25	36	100.0	471	21	US-09-791-537-92908
26	36	100.0	471	21	US-09-791-537-92908
27	36	100.0	471	21	US-09-791-537-92908
28	36	100.0	471	21	US-09-791-537-92908
29	36	100.0	471	21	US-09-791-537-92908
30	36	100.0	471	21	US-09-791-537-92908
31	36	100.0	471	21	US-09-791-537-92908
32	36	100.0	471	21	US-09-791-537-92908
33	36	100.0	471	21	US-09-791-537-92908
34	36	100.0	471	21	US-09-791-537-92908
35	36	100.0	471	21	US-09-791-537-92908
36	36	100.0	471	21	US-09-791-537-92908
37	36	100.0	471	21	US-09-791-537-92908
38	36	100.0	471	21	US-09-791-537-92908
39	36	100.0	471	21	US-09-791-537-92908
40	36	100.0	471	21	US-09-791-537-92908
41	36	100.0	471	21	US-09-791-537-92908
42	36	100.0	471	21	US-09-791-537-92908
43	36	100.0	471	21	US-09-791-537-92908
44	36	100.0	471	21	US-09-791-537-92908
45	36	100.0	471	21	US-09-791-537-92908

ALIGNMENTS

RESULT 1
US-09-791-537-94919
Sequence 94919, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791, 537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 94919
LENGTH: 448
TYPE: PRT
ORGANISM: Gallus gallus
US-09-791-537-94919

Query Match 100.0%, Score 36, DB 21, Length 448;
Best Local Similarity 100.0%, Pred. No. 3.2e+02;
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY 1 LKRVHEE 7
DB 222 LKRVHEE 228

```
RESULT 2
US-10-108-260A-4090
; Sequence 4090, Application US/10108260A
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4090
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4090

Query Match
Best Local Similarity 100.0%; Score 36; DB 25; Length 450;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
Db 219 LKKVHEE 225

RESULT 3
US-09-791-537-32570
; Sequence 32570, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32570
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Scyllorhinus stellaris
US-09-791-537-32570

Query Match
Best Local Similarity 100.0%; Score 36; DB 21; Length 455;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
Db 223 LKKVHEE 229

RESULT 4
US-09-791-537-94920
; Sequence 94920, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 94920
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Bos taurus
```

```
US-09-791-537-94920

Query Match
Best Local Similarity 100.0%; Score 36; DB 21; Length 461;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
Db 230 LKKVHEE 236

RESULT 5
US-09-791-537-94921
; Sequence 94921, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY I
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 94921
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-791-537-94921

Query Match
Best Local Similarity 100.0%; Score 36; DB 21; Length 462;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
Db 231 LKKVHEE 237

RESULT 6
US-09-791-537-73564
; Sequence 73564, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY I
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 73564
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-791-537-73564

Query Match
Best Local Similarity 100.0%; Score 36; DB 21; Length 463;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
Db 230 LKKVHEE 236

RESULT 7
US-09-538-092-948
; Sequence 948, Application US/09538092
; GENERAL INFORMATION:
```

```
;; APPLICANT: Glot, Loic
;; APPLICANT: Mansfield, Traci A.
;; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
;; FILE REFERENCE: 15966-542
;; CURRENT APPLICATION NUMBER: US/09/538,092
;; CURRENT FILING DATE: 2000-03-29
;; PRIOR APPLICATION NUMBER: 60/127,352
;; PRIOR FILING DATE: 1999-04-01
;; PRIOR APPLICATION NUMBER: 60/178,965
;; PRIOR FILING DATE: 2000-02-01
;; NUMBER OF SEQ ID NOS: 1387
;; SOFTWARE: Curataseq/formatter version 0.9
;; SEQ ID NO 948
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (0)...(0)
;; OTHER INFORMATION: Polypeptide Accession Number P17661
US-09-538-092-948

Query Match      100.0%; Score 36; DB 19; Length 469;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LKKVHEE 7
        |||||||
DB      238 LKKVHEE 244

RESULT 8
US-09-791-537-62363
;; Sequence 62363, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomimix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 62363
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Rattus norvegicus
US-09-791-537-62363

Query Match      100.0%; Score 36; DB 21; Length 469;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LKKVHEE 7
        |||||||
DB      238 LKKVHEE 244

RESULT 9
US-09-791-537-73566
;; Sequence 73566, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomimix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
```

```
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 73566
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Mesocricetus auratus
US-09-791-537-73566

Query Match      100.0%; Score 36; DB 21; Length 469;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LKKVHEE 7
        |||||||
DB      238 LKKVHEE 244

RESULT 10
US-09-791-537-132613
;; Sequence 132613, Application US/09791537
;; GENERAL INFORMATION:
;; APPLICANT: Biomimix, Inc.
;; APPLICANT: Debe, Derek
;; APPLICANT: Danzer, Joseph
;; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: 261/210
;; CURRENT APPLICATION NUMBER: US/09/791,537
;; CURRENT FILING DATE: 2001-02-22
;; NUMBER OF SEQ ID NOS: 153055
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 132613
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-09-791-537-132613

Query Match      100.0%; Score 36; DB 21; Length 469;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LKKVHEE 7
        |||||||
DB      238 LKKVHEE 244

RESULT 11
US-60-389-987-213
;; Sequence 213, Application US/60389987
;; GENERAL INFORMATION:
;; APPLICANT: Ghosh, Soumitra S.
;; APPLICANT: Fahy, Eoin D.
;; APPLICANT: Zhang, Bing
;; APPLICANT: Gibson, Bradford W.
;; APPLICANT: Taylor, Steven W.
;; APPLICANT: Glenn, Gary M.
;; APPLICANT: Warnock, Dale E.
;; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
;; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
;; FILE REFERENCE: 660088,465P2
;; CURRENT APPLICATION NUMBER: US/60/389,987
;; CURRENT FILING DATE: 2002-06-17
;; NUMBER OF SEQ ID NOS: 3025
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 213
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-389-987-213

Query Match      100.0%; Score 36; DB 27; Length 469;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 LKKVHEE 7
|||||
DB 238 LKKVHEE 244

RESULT 12
US-60-412-418-213
; Sequence 213, Application US/60412418
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojn D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465P3
; CURRENT APPLICATION NUMBER: US/60/412,418
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 3025
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 213
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-412-418-213

Query Match 100.0%; Score 36; DB 27; Length 469;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
|||||
DB 238 LKKVHEE 244

RESULT 13
US-09-791-537-46962
; Sequence 46962, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 46962
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-791-537-46962

Query Match 100.0%; Score 36; DB 21; Length 470;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
|||||
DB 239 LKKVHEE 245

RESULT 14
US-09-791-537-62106
; Sequence 62106, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blomomix, Inc.
; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 62106
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-62106

Query Match 100.0%; Score 36; DB 21; Length 470;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
|||||
DB 239 LKKVHEE 245

RESULT 15
US-09-791-537-65178
; Sequence 65178, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Blomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 65178
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-65178

Query Match 100.0%; Score 36; DB 21; Length 470;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
|||||
DB 239 LKKVHEE 245

Search completed: November 13, 2002, 13:39:15
Job time : 142.447 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:17:03 ; Search time 2.82979 Seconds
(without alignments)
29.431 Million cell updates/sec

Title: US-09-856-086-2
Perfect score: 36
Sequence: 1 LKRVHE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 41632 seqs, 11897606 residues

Total number of hits satisfying chosen parameters: 41632

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	7	5	US-09-856-086-2
2	32	86.9	6	5	US-09-856-086-5
3	31	86.1	76	5	US-09-513-999C-7614
4	31	86.1	600	5	US-10-264-237-1886
5	30	83.3	804	5	US-10-092-411A-5218
6	28	77.8	66	5	US-09-513-999C-5468
7	28	77.8	135	6	US-10-141-531-23
8	28	77.8	154	6	US-10-092-411A-4529
9	28	77.8	280	6	US-10-125-923A-30
10	28	77.8	315	6	US-10-141-531-66
11	28	77.8	316	6	US-10-141-531-85
12	28	77.8	316	6	US-10-141-531-94
13	28	77.8	316	6	US-10-141-531-103
14	28	77.8	316	6	US-10-141-531-112
15	28	77.8	316	6	US-10-141-531-121
16	28	77.8	316	6	US-10-141-531-130
17	28	77.8	316	6	US-10-141-531-139
18	28	77.8	316	6	US-10-141-531-148
19	28	77.8	316	6	US-10-141-531-157
20	28	77.8	316	6	US-10-141-531-166
21	28	77.8	316	6	US-10-141-531-175
22	28	77.8	316	6	US-10-141-531-184
23	28	77.8	316	6	US-10-141-531-193
24	28	77.8	316	6	US-10-141-531-202
25	28	77.8	581	6	US-10-085-198-122
26	28	77.8	619	6	US-10-092-411A-4248

27	28	77.8	988	6	US-10-167-631A-3	Sequence 3, Appl1
28	28	77.8	1017	6	US-10-167-631A-4	Sequence 4, Appl1
29	28	77.8	1707	6	US-10-167-631A-2	Sequence 2, Appl1
30	27	75.0	61	5	US-09-513-999C-7340	Sequence 7340, Ap
31	27	75.0	105	6	US-10-092-411A-5207	Sequence 5207, Ap
32	27	75.0	146	6	US-10-092-411A-2943	Sequence 2943, Ap
33	27	75.0	234	6	US-10-154-678-20	Sequence 20, Appl1
34	27	75.0	248	5	US-09-849-772-8	Sequence 8, Appl1
35	27	75.0	255	5	US-09-849-772-4	Sequence 4, Appl1
36	27	75.0	561	6	US-10-136-728-70	Sequence 70, Appl1
37	27	75.0	561	6	US-10-136-728-72	Sequence 72, Appl1
38	26	72.2	100	5	US-09-513-999C-4752	Sequence 4752, Ap
39	26	72.2	151	6	US-10-264-237-2018	Sequence 2018, Ap
40	26	72.2	307	6	US-10-092-411A-5144	Sequence 5144, Ap
41	26	72.2	372	6	US-10-094-507-35	Sequence 35, Appl1
42	26	72.2	443	6	US-10-113-709A-2	Sequence 2, Appl1
43	26	72.2	538	6	US-10-092-411A-3463	Sequence 3463, Ap
44	26	72.2	694	6	US-10-125-923A-60	Sequence 60, Appl1
45	26	72.2	823	6	US-10-092-411A-4081	Sequence 4081, Ap

ALIGNMENTS

RESULT 1
US-09-856-086-2
; Sequence 2, Application US/09856086
; GENERAL INFORMATION:
; APPLICANT: EBRINGER, ALAN
; TITLE OF INVENTION: DIAGNOSIS OF DEMENTIATING OR SPONGIFORM DISEASE
; FILE REFERENCE: 78104.040
; CURRENT APPLICATION NUMBER: US/09/856,086
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens, Bos
US-09-856-086-2

Query Match 100.0%; Score 36; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.8e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHE 7
DB 1 LKRVHE 7

RESULT 2
US-09-856-086-5
; Sequence 5, Application US/09856086
; GENERAL INFORMATION:
; APPLICANT: EBRINGER, ALAN
; TITLE OF INVENTION: DIAGNOSIS OF DEMENTIATING OR SPONGIFORM DISEASE
; FILE REFERENCE: 78104.040
; CURRENT APPLICATION NUMBER: US/09/856,086
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens, Bos
US-09-856-086-5

Query Match 88.9%; Score 32; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 3.8e+04;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LKRVHE 7
|||||

Db 1 KKVHHE 6

RESULT 3
US-09-513-999C-7614Sequence 7614, Application US/09513999C
GENERAL INFORMATION:
APPLICANT: Dumas Mline Edwards, J.B.APPLICANT: Duclet, A.
APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

FILE REFERENCE: 59. US2. REG

CURRENT APPLICATION NUMBER: US/09/513, 999C

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/122,487

NUMBER OF SEQ ID NOS: 36681

SOFTWARE: Patent. pm

SEQ ID NO 7614

LENGTH: 76

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: UNSURELOCATION: 3
OTHER INFORMATION: Xaa-Ala or GlyFEATURE:
NAME/KEY: UNSURELOCATION: 4
OTHER INFORMATION: Xaa-Ile or MetFEATURE:
NAME/KEY: UNSURELOCATION: 5
OTHER INFORMATION: Xaa-Asp or AsnFEATURE:
NAME/KEY: UNSURELOCATION: 9
OTHER INFORMATION: Xaa-Phe or ValFEATURE:
NAME/KEY: UNSURELOCATION: 46
OTHER INFORMATION: Xaa-Cys or Phe or His or Ile or Leu or Asn or Pro or Arg or Ser

US-09-513-999C-7614

Query Match 86.1%; Score 31; DB 5; Length 76;
Best Local Similarity 71.4%; Pred. No. 4.8;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHHE 7

Db 61 IKKLHEE 67

RESULT 4
US-10-264-237-1886Sequence 1886, Application US/10264237
GENERAL INFORMATION:
APPLICANT: Birse et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: P131PI

CURRENT APPLICATION NUMBER: US/10/264,237

PRIOR FILING DATE: 2002-10-04

PRIOR APPLICATION NUMBER: PCT/US01/16450

PRIOR FILING DATE: 2001-05-18

PRIOR APPLICATION NUMBER: US 60/205,515

NUMBER OF SEQ ID NOS: 2876

SOFTWARE: Patent Ver. 3.1

SEQ ID NO 1886

LENGTH: 600

TYPE: PRT
ORGANISM: Homo sapiens

US-10-264-237-1886

Query Match 86.1%; Score 31; DB 6; Length 600;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHE 6

Db 222 LKKVHE 227

RESULT 5
US-10-092-411A-5218Sequence 5218, Application US/10092411A
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCO

FILE REFERENCE: 032796-101

CURRENT APPLICATION NUMBER: US/10/092,411A

CURRENT FILING DATE: 2002-03-07

PRIOR APPLICATION NUMBER: US 09/134,001

PRIOR FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08

PRIOR APPLICATION NUMBER: US 60/055,779

NUMBER OF SEQ ID NOS: 5676

SEQ ID NO 5218

LENGTH: 804

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-10-092-411A-5218

Query Match 83.3%; Score 30; DB 6; Length 804;
Best Local Similarity 85.7%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LKKVHEE 7

Db 440 LKKEHEE 446

RESULT 6
US-09-513-999C-5468Sequence 5468, Application US/09513999C
GENERAL INFORMATION:
APPLICANT: Dumas Mline Edwards, J.B.

APPLICANT: Duclet, A.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

FILE REFERENCE: 59. US2. REG

CURRENT APPLICATION NUMBER: US/09/513, 999C

CURRENT FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/122,487

PRIOR FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 36681

SOFTWARE: Patent. pm

SEQ ID NO 5468

LENGTH: 66

TYPE: PRT

ORGANISM: Homo sapiens

US-09-513-999C-5468

Query Match 77.8%; Score 28; DB 5; Length 66;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LKKVHEE 7

Db 14 LKGIHEE 20

RESULT 7

```

US-10-141-531-23
; Sequence 23, Application US/10141531
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Dalma, Bipin K.
; APPLICANT: del Val, Greg
; APPLICANT: Desjarlais, John R.
; APPLICANT: Helfetz, Peter
; APPLICANT: Lugibuhl, Peter
; APPLICANT: Muchhal, Umesh
; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
; FILE REFERENCE: A-71457-2/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/141,531
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US 60/370,609
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/289,029
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-141-531-23

```

```

Query Match          77.8%; Score 28; DB 6; Length 135;
Best Local Similarity 57.1%; Pred. No. 32;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LKRYHEE 7
Db 23 VKRIHEE 29

```

```

RESULT 8
US-10-092-411A-4529
; Sequence 4529, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 4529
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-4529

```

```

Query Match          77.8%; Score 28; DB 6; Length 154;
Best Local Similarity 71.4%; Pred. No. 37;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LKRYHEE 7
Db 87 LKRIHEE 93

```

```

RESULT 9
US-10-125-923A-30
; Sequence 30, Application US/10125923A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian

```

```

; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zhen
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C79
; CURRENT APPLICATION NUMBER: US/10/125,923A
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 30
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-125-923A-30

```

```

Query Match          77.8%; Score 28; DB 6; Length 280;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LKRYHEE 7
Db 232 LKRYHEE 238

```

```

RESULT 10
US-10-141-531-66
; Sequence 66, Application US/10141531
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Dalma, Bipin K.
; APPLICANT: del Val, Greg
; APPLICANT: Desjarlais, John R.
; APPLICANT: Helfetz, Peter
; APPLICANT: Lugibuhl, Peter
; APPLICANT: Muchhal, Umesh
; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
; FILE REFERENCE: A-71457-2/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/141,531
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US 60/370,609
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/289,029
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66

```

LENGTH: 315
TYPE: PRT
ORGANISM: Bacillus subtilis
US-10-141-531-66

Query Match
Best Local Similarity 77.8%; Score 28; DB 6; Length 315;
Pred. No. 74;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKKVHEE 7
:|:|:|
Db 203 VKEIHEE 209

RESULT 11
US-10-141-531-85
Sequence 85, Application US/10141531
GENERAL INFORMATION:
APPLICANT: Briggs, Steven P.
APPLICANT: Dalmla, Bipin K.
APPLICANT: del Val, Greg
APPLICANT: Desjarlais, John R.
APPLICANT: Helfetz, Peter
APPLICANT: Lugnbuhl, Peter
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
FILE REFERENCE: A-71457-2/RT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/141,531
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: US 60/370,609
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: US 60/289,029
PRIOR FILING DATE: 2001-05-04
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PatentIn version 3.1
SEQ ID NO 85
LENGTH: 316
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-85

Query Match
Best Local Similarity 77.8%; Score 28; DB 6; Length 316;
Pred. No. 74;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKKVHEE 7
:|:|:|
Db 204 VKEIHEE 210

RESULT 12
US-10-141-531-94
Sequence 94, Application US/10141531
GENERAL INFORMATION:
APPLICANT: Briggs, Steven P.
APPLICANT: Dalmla, Bipin K.
APPLICANT: del Val, Greg
APPLICANT: Desjarlais, John R.
APPLICANT: Helfetz, Peter
APPLICANT: Lugnbuhl, Peter
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
FILE REFERENCE: A-71457-2/RT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/141,531
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: US 60/370,609
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: US 60/289,029
PRIOR FILING DATE: 2001-05-04
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PatentIn version 3.1

SEQ ID NO 94
LENGTH: 316
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-94

Query Match
Best Local Similarity 77.8%; Score 28; DB 6; Length 316;
Pred. No. 74;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKKVHEE 7
:|:|:|
Db 204 VKEIHEE 210

RESULT 13
US-10-141-531-103
Sequence 103, Application US/10141531
GENERAL INFORMATION:
APPLICANT: Briggs, Steven P.
APPLICANT: Dalmla, Bipin K.
APPLICANT: del Val, Greg
APPLICANT: Desjarlais, John R.
APPLICANT: Helfetz, Peter
APPLICANT: Lugnbuhl, Peter
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
FILE REFERENCE: A-71457-2/RT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/141,531
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: US 60/370,609
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: US 60/289,029
PRIOR FILING DATE: 2001-05-04
NUMBER OF SEQ ID NOS: 208
SOFTWARE: PatentIn version 3.1
SEQ ID NO 103
LENGTH: 316
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-103

Query Match
Best Local Similarity 77.8%; Score 28; DB 6; Length 316;
Pred. No. 74;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKKVHEE 7
:|:|:|
Db 204 VKEIHEE 210

RESULT 14
US-10-141-531-112
Sequence 112, Application US/10141531
GENERAL INFORMATION:
APPLICANT: Briggs, Steven P.
APPLICANT: Dalmla, Bipin K.
APPLICANT: del Val, Greg
APPLICANT: Desjarlais, John R.
APPLICANT: Helfetz, Peter
APPLICANT: Lugnbuhl, Peter
TITLE OF INVENTION: Nucleic Acids and Proteins with Thioredoxin Reductase Activity
FILE REFERENCE: A-71457-2/RT/RMS/RMK
CURRENT APPLICATION NUMBER: US/10/141,531
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION NUMBER: US 60/370,609
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: US 60/289,029

; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 112
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-112

Query Match 77.88; Score 28; DB 6; Length 316;
Best Local Similarity 57.18; Pred. No. 74;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKXVHEE 7
Db 204 VKXIHHEE 210

RESULT 15
US-10-141-531-121
; Sequence 121, Application US/10141531
; GENERAL INFORMATION:
; APPLICANT: Briggs, Steven P.
; APPLICANT: Daimla, Bipin K.
; APPLICANT: del Val, Greg
; APPLICANT: Desjarlais, John R.
; APPLICANT: Heifetz, Peter
; APPLICANT: Luginduhl, Peter
; APPLICANT: Muchhal, Umesh
; TITLE OF INVENTION: Nucleic Acids and Proteins with Thioedoxin Reductase Activity
; FILE REFERENCE: A-71457-2/REF/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/141,531
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: US 60/370,609
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/289,029
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 121
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Bacillus subtilis variant
US-10-141-531-121

Query Match 77.88; Score 28; DB 6; Length 316;
Best Local Similarity 57.18; Pred. No. 74;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKXVHEE 7
Db 204 VKXIHHEE 210

Search completed: November 13, 2002, 13:39:38
Job time : 2.82979 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2002, 13:13:25 ; Search time 10.8723 Seconds
(without alignments)
61.895 Million cell updates/sec

Title: US-09-856-086-2
Perfect score: 36
Sequence: 1 LKKVHEE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	36	100.0	287 2 A21762	neurofilament trip
2	36	100.0	298 1 DMHY	desmin - golden ha
3	36	100.0	463 1 DMCH	desmin - chicken
4	36	100.0	469 1 DMHU	desmin - human
5	36	100.0	469 2 A52469	desmin - rat
6	36	100.0	469 2 A24783	desmin - golden ha
7	36	100.0	469 2 A54104	desmin - mouse
8	36	100.0	543 1 OPMSL	neurofilament trip
9	36	100.0	544 2 S07144	neurofilament trip
10	36	100.0	544 2 B44841	neurofilament trip
11	36	100.0	548 1 OPEGL	neurofilament trip
12	36	100.0	554 2 JF0094	neurofilament trip
13	36	100.0	556 2 B46024	neurofilament-li su
14	35	97.2	410 2 T19995	hypothetical prote
15	35	97.2	458 2 A43534	desmin - African c
16	35	97.2	807 2 T00990	hypothetical prote
17	34	94.4	68 2 A49830	ventricular myosin
18	33	91.7	424 2 S14887	peripherin (clone
19	33	91.7	438 2 S07823	peripherin (clone
20	33	91.7	456 2 A60090	peripherin - Afric
21	33	91.7	456 2 S08228	intermediate filam
22	33	91.7	466 2 A25074	vimentin - human
23	33	91.7	468 2 JN0016	peripherin interne
24	33	91.7	469 2 A44841	low molecular weigh
25	33	91.7	471 2 A55185	intermediate filam
26	33	91.7	823 2 B81282	probable integral
27	33	91.7	931 2 G68615	ATP-dependent DNA
28	33	91.7	3135 2 A48584	transmission block
29	32	88.9	30 2 S74121	fructose-bisphosph

30	32	88.9	144 2 E96618	hypothetical prote
31	32	88.9	189 2 S39864	late competence op
32	32	88.9	243 2 B97098	uncharacterized co
33	32	88.9	308 2 D98095	conserved hypothet
34	32	88.9	317 2 T42645	hypothetical prote
35	32	88.9	400 2 F64613	hypothetical prote
36	32	88.9	400 2 G71900	hypothetical prote
37	32	88.9	430 2 I56572	glial fibrillary a
38	32	88.9	432 2 A32936	glial fibrillary a
39	31	86.1	134 2 T31726	hypothetical prote
40	31	86.1	159 2 AD2691	conserved hypothet
41	31	86.1	162 2 H97472	hypothetical prote
42	31	86.1	173 2 AG3340	4-hydroxyphenylace
43	31	86.1	328 2 D72393	oligopeptide ABC t
44	31	86.1	329 2 F64356	translation initia
45	31	86.1	333 2 F70341	sulfur oxygenase r

ALIGNMENTS

RESULT 1
A21762
neurofilament triplet L protein - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 13-Aug-1999
C:Accession: A21762
R:Julien, J.P.; Ramchandran, K.; Grosfeld, F.
Biochim. Biophys. Acta 825, 398-404, 1985
A:Title: Cloning of a cDNA encoding the smallest neurofilament protein from the rat.
A:Reference number: A21762; PMID:85252830; PMID:3925999
A:Accession: A21762
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-287 <TUL>
A:Cross-References: GB:M25638; NID:9205683; PIDN:AAA1694.1; PID:9205684
A:Superfamily: cytoskeletal Keratin
C:Keywords: coiled coil

Query Match
Best Local Similarity 100.0%; Score 36; DB 2; Length 287;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
DB 26 LKKVHEE 32

RESULT 2
DMHY
desmin - golden hamster (fragment)
C:Species: Mesocricetus auratus (golden hamster)
C>Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 22-Jun-1999
C:Accession: A02956
R:Quax, W.; van den Heuvel, R.; Egberts, W.V.; Quax-Jeukens, Y.; Bloemendaal, H.
Proc. Natl. Acad. Sci. U.S.A. 81, 5970-5974, 1984
A:Title: Intermediate filament cDNAs from BHK-21 cells: demonstration of distinct gen
A:Reference number: A02956; PMID:85014850; PMID:6091127
A:Accession: A02956
A:Molecule type: mRNA
A:Residues: 1-298 <QUA>
A:Cross-References: GB:K02407; NID:9191356; PIDN:AAA37071.1; PID:9387070
A:Experimental source: baby hamster kidney cells, BHK-21
C:Comment: There is a single gene for desmin in the hamster genome.
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil; intermediate filament
F:245-298/Domain: tail <TLE>

Query Match
Best Local Similarity 100.0%; Score 36; DB 1; Length 298;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7

Db 67 LKVVHEE 73

RESULT 3

desmin - chicken

N:Alternate names: type III intermediate filament

C:Species: Gallus gallus (chicken)

C:Date: 18-Dec-1981 #sequence_revision 12-Apr-1996 #text_change 10-Dec-1999

C:Accession: A00969; A94014; J01459; S02448; A32858; S23189; A02957

R:Geisler, N.; Weber, K.

EMBO J. 1, 1649-1656, 1982

A:Title: The amino acid sequence of chicken muscle desmin provides a common structural

A:Reference number: A00969; PMID:84207925; PMID:6202512

A:Accession: A00969

A:Molecule type: protein

A:Residues: 1-463 <GEI>

R:Capetanaki, Y.G.; Ngai, J.; Lazarides, E.

Proc. Natl. Acad. Sci. U.S.A. 81, 6909-6913, 1984

A:Title: Characterization and regulation in the expression of a gene coding for the inte

A:Reference number: A94014; PMID:85063701; PMID:6594672

A:Accession: A94014

A:Molecule type: mRNA

A:Residues: 366-463 <CAP>

A:Cross-references: GB:K02445; NID:9211727; PID:AAA48751.1; PID:9211728

R:Kusubata, M.; Matsuno, Y.; Tsujimura, K.; Ito, H.; Ando, S.; Kamijo, M.; Yasuda, H.

Biochem. Biophys. Res. Commun. 190, 927-934, 1993

A:Title: cdc2 kinase phosphorylation of desmin at three serine/threonine residues in the

A:Reference number: J01459; PMID:93176201; PMID:8439342

A:Accession: J01459

A:Molecule type: protein

A:Residues: 5-9;16-27;63-67 <KUS>

R:Geisler, N.; Weber, K.

EMBO J. 7, 15-20, 1988

A:Title: Phosphorylation of desmin in vitro inhibits formation of intermediate filaments

A:Reference number: S02448; PMID:88196075; PMID:3359992

A:Accession: S02448

A:Molecule type: protein

A:Residues: 1-69 <GEI>

R:Kikunaga, S.; Ando, S.; Shibata, M.; Tanabe, K.; Sato, C.; Inagaki, M.

J. Biol. Chem. 264, 5674-5678, 1989

A:Title: Protein kinase C phosphorylation of desmin at four serine residues within the r

A:Reference number: A32858; PMID:89174618; PMID:2494166

A:Accession: A32858

A:Molecule type: protein

A:Residues: 10-14;28-42;49-59 <KIT>

R:Geisler, N.; Schuenemann, J.; Weber, K.

Eur. J. Biochem. 206, 841-852, 1992

A:Title: Chemical cross-linking indicates a staggered and antiparallel protofilament of

A:Reference number: S23189; PMID:92299013; PMID:1606966

A:Accession: S23189

A:Status: preliminary

A:Molecule type: protein

A:Residues: 110-118;255-266;274-282;393-401 <GEI>

C:Comment: This protein was isolated from chicken gizzard.

C:Comment: There appears to be a single desmin gene in the haploid chicken genome that i

C:Comment: Desmin intermediate filaments are found in the cytoplasm of cultured myogeni

C:Comment: The molecule contains three structurally distinct domains. The surface-exposed

C:Superfamily: cytoskeletal keratin

C:Keywords: blocked amino end; coiled coil; intermediate filament; muscle; phosphoprotei

F:1-99/Domain: head <HEI>

F:100-407/Domain: rod <ROD>

F:100-132/Region: coil 1A

F:133-142/Region: linker 1

F:143-243/Region: coil 1B

F:244-259/Region: linker 12

F:260-278/Region: linker 2A

F:279-286/Region: linker 2

F:287-407/Region: coil 2B

F:345/Region: stutter

F:408-463/Domain: tail <END>

F:5/22/Binding site: blocked amino end (Ser) (probably acetylated) #status experiment

F:12/29/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #status experiment

F:29/35/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status

F:64/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status experimental

Query Match

Best Local Similarity 100.0%; Score 36; DB 1; Length 463;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKVVHEE 7

Db 230 LKVVHEE 236

RESULT 4

desmin - human

N:Alternate names: type III intermediate filament

C:Species: Homo sapiens (man)

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 10-Dec-1999

C:Accession: J00063

R:L.L. Lillienbaum, A.; Butler-Browne, G.; Paulin, D.

Gene 78, 243-254, 1989

A:Title: Human desmin-coding gene: complete nucleotide sequence, characterization a

A:Reference number: J00063; PMID:89378751; PMID:2673923

A:Accession: J00063

A:Molecule type: DNA

A:Residues: 1-469 <LIT>

A:Cross-references: GB:M63391; GB:M26935; GB:M58168; GB:M59379; GB:M65071; GB:M5315

A:Note: The introns of this gene contain 1.2K of repetitive sequences belonging to

A:Comment: Desmin intermediate filaments are found in the cytoplasm of cultured myo

A:Comment: The molecule contains three structurally distinct domains. The surface-e

A:Comment: The tailpiece comprises the carboxyl-terminal residues.

C:Genetics:

A:Gene: GDB:DES

A:Cross-references: GDB:119841; OMIM:125660

A:Map position: 2q35-2q35

A:Introns: 192/2; 212/3; 244/3; 298/3; 340/3; 414/2; 429/1; 456/3

C:Superfamily: cytoskeletal keratin

C:Keywords: acetylated amino end; coiled coil; intermediate filament; muscle; phosp

F:2-469/Product: desmin #status predicted <MDM>

F:2-108/Domain: head <HEI>

F:109-415/Domain: rod <ROD>

F:416-469/Domain: tail <END>

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted

F:7/32/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #status predicted

F:13/48/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pred

F:45/60/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status

Query Match

Best Local Similarity 100.0%; Score 36; DB 1; Length 469;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKVVHEE 7

Db 238 LKVVHEE 244

RESULT 5

desmin - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 13-Aug-1999

C:Accession: I52469

R:van Groningen, J.U.; Bloemers, H.P.; Swart, G.W.

Biochim. Biophys. Acta 1217, 107-109, 1994

A:Title: Rat desmin gene structure and expression.

A:Reference number: I52469; PMID:94114566; PMID:8286410

A:Accession: I52469

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-469 <RES>
 A:Cross-references: EMBL:X73524; NID:9452778; PIDN:CA51920.1; PID:9452779
 C:Superfamily: cytoskeletal keratin

Query Match 100.0%; Score 36; DB 2; Length 469;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
 |||||
 Db 238 LKKVHEE 244

RESULT 6

A24783
 desmin - golden hamster
 C:Species: Mesocricetus auratus (golden hamster)
 C:Date: 22-Jul-1987 #sequence_revision 22-Jul-1987 #text_change 13-Aug-1999
 A:Accession: A24783
 R:Quax, W.; van den Broek, L.; Egberts, W.V.; Ramaekers, F.; Bloemendal, H.
 Cell 43, 327-338, 1985
 A:Title: Characterization of the hamster desmin gene: expression and formation of desmin
 A:Reference number: A24783; MUID:86079506; PMID:3855248
 A:Accession: A24783
 A:Molecule type: DNA
 A:Residues: 1-469 <QU>
 A:Cross-references: GB:M12104; NID:9191360; PIDN:AAA37072.1; PID:9387071
 C:Genetics:
 A:Introns: 192/2; 212/3; 244/3; 298/3; 340/3; 414/2; 429/1; 456/3
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil

Query Match 100.0%; Score 36; DB 2; Length 469;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
 |||||
 Db 238 LKKVHEE 244

RESULT 7

A54104
 desmin - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 19-Apr-1996
 A:Accession: A54104; S31404
 R:Li, H.; Choudhary, S.K.; Milner, D.J.; Munir, M.I.; Kulsk, I.R.; Capetanaki, Y.
 J. Cell Biol. 124, 827-841, 1994
 A:Title: Inhibition of desmin expression blocks myoblast fusion and interferes with the
 A:Reference number: A54104; MUID:94165148; PMID:8120103
 A:Accession: A54104
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-469 <LIA>
 A:Cross-references: GB:L22550
 R:Li, H.; Capetanaki, Y.
 Submitted to the EMBL Data Library, December 1992
 A:Description: Regulation of the Mouse Desmin Gene: Transactivation by MyoD, Myogenin, M
 A:Reference number: S31404
 A:Accession: S31404
 A:Molecule type: DNA
 A:Residues: 1-40 <LIH>
 A:Cross-references: EMBL:218692
 C:Superfamily: cytoskeletal keratin
 C:Keywords: muscle

Query Match 100.0%; Score 36; DB 2; Length 469;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
 |||||
 Db 238 LKKVHEE 244

RESULT 8

QFMSL
 neurofilament triplet L protein - mouse
 N:Alternate names: 68K neurofilament protein; NF-L(low) protein; type IV IF protein
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Mar-1988 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999
 A:Accession: A25227; A26562; A43772; A41012; I55316
 R:Lewis, S.A.; Cowan, N.J.
 Mol. Cell. Biol. 6, 1529-1534, 1986
 A:Title: Anomalous placement of introns in a member of the intermediate filament mult
 A:Reference number: A25227; MUID:87064433; PMID:3785173
 A:Accession: A25227
 A:Molecule type: DNA
 A:Residues: 1-543 <LEM>
 A:Cross-references: GB:M13016; NID:9200023; PIDN:AAA39810.1; PID:9387492
 A:Note: the authors translated the codon GGC for residue 5 as Ala, ACA for residue 88
 1 as Glu

R:Lewis, S.A.; Cowan, N.J.
 J. Cell Biol. 100, 843-850, 1985
 A:Title: Genetics, evolution, and expression of the 68,000-mol-wt neurofilament prote
 A:Reference number: A26562; MUID:85131334; PMID:3919033
 A:Accession: A26562
 A:Molecule type: mRNA
 A:Residues: 242-543 <LE2>
 A:Cross-references: GB:X02165
 A:Experimental source: Brain
 R:Julien, J.P.; Meyer, D.; Flavell, D.; Hurst, J.; Grosfeld, F.
 Brain Res. Mol. Brain Res. 1, 243-250, 1986
 A:Title: Cloning and developmental expression of the murine neurofilament gene family
 A:Reference number: A43772
 A:Accession: A43772
 A:Molecule type: mRNA
 A:Residues: 1-5, 'Y', '7-8, 'Y', '10-64, 'W', '66-72, 'L', '74-98, 'D', '100-194, 'R', '196-202, 204-239
 A:Cross-references: GB:M20480; NID:9200037; PIDN:AAA39814.1; PID:9200038
 A:Note: the authors translated the codon CCG for residue 195 as Ala
 R:Singh, R.K.; Nixon, R.A.
 J. Biol. Chem. 266, 18861-18867, 1991

A:Title: Identification of Ser-55 as a major protein kinase A phosphorylation site on
 A:Reference number: A41012; MUID:92011653; PMID:1717455
 A:Accession: A41012
 A:Molecule type: protein
 A:Residues: 52-57 <SIH>
 R:Nakahira, K.; Ikemura, K.; Wade, K.; Tamura, T.
 J. Biol. Chem. 265, 19786-19791, 1990
 A:Title: Structure of the 68-kDa neurofilament gene and regulation of its expression.
 A:Reference number: I55316; MUID:91060592; PMID:2246261
 A:Accession: I55316
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-5, 'Y', '7-8, 'Y', '10-28 <RES>
 A:Cross-references: GB:M55423; NID:9200027; PIDN:AAA39812.1; PID:9554245
 C:Comment: This is the most abundant of the three neurofilament proteins and, as the
 C:Genetics:
 A:Introns: 349/3; 391/2; 498/1
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:2-72/Domain: head <HED>
 F:94-125/Domain: coil 1a, alpha-helical rod #status predicted <RIA>
 F:126-138/Domain: linker 1
 F:139-234/Domain: coil 1b, alpha-helical rod #status predicted <RIb>
 F:235-236/Domain: linker 12
 F:257-272/Domain: coil 2a, alpha-helical rod #status predicted <R2a>
 F:273-281/Domain: linker 2
 F:282-401/Domain: coil 2b, alpha-helical rod #status predicted <R2b>
 F:404-543/Domain: tail <TAI>
 F:404-444/Region: tail subdomain a
 F:445-543/Region: tail subdomain b

Query Match 100.0%; Score 36; DB 1; Length 543;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
 |||||
 Db 223 LKRVHEE 229

RESULT 9

S07144
 neurofilament triplet L protein - human
 N:Alternate names: neurofilament light polypeptide (68K)
 N:Contains: Glu-50 brain peptide
 C:Species: Homo sapiens (man)
 C>Date: 29-Jan-1993 #sequence, revision 29-Jan-1993 #text_change 21-Jul-2000
 C:Accession: S07144; 152832; A60703

R:Jullien, J.P.; Grosfeld, F.; Yazdankhah, K.; Flavell, D.; Meijer, D.; Mushynski, W.
 Blochim. Biophys. Acta 909, 10-20, 1987
 A:Title: The structure of a human neurofilament gene (NF-L): a unique exon-intron organ
 A:Reference number: S07144; MUID:87214213; PMID:3034332
 A:Accession: S07144

A:Molecule type: DNA
 A:Residues: 1-544 <TUL>
 A:Cross-references: EMBL:X05608; NID:94195072; PIDN:CA429097.1; PID:941279504
 A:Note: the authors translated the codon ATG for residue 366 as Asn
 R:Pospelov, V.A.; Pospelova, T.V.; Jullien, J.P.
 Cell Growth Differ. 5, 187-196, 1994
 A:Title: AP-1 and Krox-24 transcription factors activate the neurofilament light gene pr
 A:Reference number: 152832; MUID:94235564; PMID:8180132
 A:Accession: 152832

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-10 <POS>
 A:Cross-references: GB:S70309; NID:9547176; PIDN:AAD14057.1; PID:94261757
 R:Nomata, Y.; Matanabe, T.; Wada, H.
 J. Biochem. 93, 825-831, 1983
 A:Title: Highly acidic proteins from human brain: purification and properties of Glu-50
 A:Reference number: A60703; MUID:83265667; PMID:6135695
 A:Accession: A60703

A:Molecule type: protein
 A:Residues: 469-472, 'D', 474 <NOM>
 A:Experimental source: Glu-50 brain peptide
 A:Note: this acidic protein is named for its greater than fifty per cent glutamic acid c
 C:Genetics:
 A:Gene: GDB:NEFL; NFL
 A:Cross-references: GDB:120227; OMIM:162280
 A:Map position: 8p21-8p21
 A:Introns: 349/3; 391/2; 498/1
 C:Superfamily: cytoskeletal keratin
 C:Keywords: brain; coiled coil; intermediate filament
 F:469-544/Product: Glu-50 peptide #status predicted <ES0>

Query Match 100.0%; Score 36; DB 2; Length 544;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
 |||||
 Db 224 LKRVHEE 230

RESULT 10

B44841
 low molecular weight neurofilament protein XNF-L - African clawed frog

C:Species: Xenopus laevis (African clawed frog)
 C>Date: 31-Mar-1993 #sequence, revision 18-Nov-1994 #text_change 12-Apr-1995
 C:Accession: B44841
 R:Charnas, L.R.; Szaro, B.G.; Galner, H.
 J. Neurosci. 12, 3010-3024, 1992
 A:Title: Identification and developmental expression of a novel low molecular weight neu
 A:Reference number: A44841; MUID:92363194; PMID:1494944
 A:Accession: B44841

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-544 <CHA>
 A:Experimental source: brain
 A:Note: sequence inconsistent with the nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:110225; NCBI:110226)
 C:Superfamily: cytoskeletal keratin

Query Match 100.0%; Score 36; DB 2; Length 544;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
 |||||
 Db 217 LKRVHEE 223

RESULT 11

Q9PGL
 neurofilament triplet L protein - pig

N:Alternate names: 68K neurofilament protein
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 15-Nov-1984 #sequence, revision 28-May-1986 #text_change 10-Oct-1997
 C:Accession: A91337; A90973; A34569; A02963
 R:Geisler, N.; Plessmann, U.; Weber, K.
 FEBS Lett. 182, 475-478, 1985
 A:Title: The complete amino acid sequence of the major mammalian neurofilament prot
 A:Reference number: A91337; MUID:85154583; PMID:3920075
 A:Accession: A91337

A:Molecule type: protein
 A:Residues: 1-547 <GEI>
 R:Geisler, N.; Kaufmann, E.; Fischer, S.; Plessmann, U.; Weber, K.
 EMBO J. 2, 1295-1302, 1983
 A:Title: Neurofilament architecture combines structural principles of intermediate
 A:Reference number: A90973
 A:Accession: A90973

A:Molecule type: protein
 A:Residues: 1-82; 278-548 <GE2>
 A:Note: residue 322 is either lysine or arginine
 R:Gonda, Y.; Nishizawa, K.; Ando, S.; Kitamura, S.; Minoura, Y.; Nishit, Y.; Inagaki
 Blochim. Biophys. Res. Commun. 167, 1316-1325, 1990
 A:Title: Involvement of protein kinase C in the regulation of assembly-disassembly
 A:Reference number: A34569; MUID:90211310; PMID:2108674
 A:Accession: A34569

A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 9-14; 23-29; 30-53 <GON>
 C:Comment: Mammalian neurofilaments usually contain three polypeptides, L, M, and H
 ke all other intermediate filament proteins: a conserved alpha-helical region, whos
 al beta turns; domain b is acidic and rich in glutamic acid and lysine residues
 C:Comment: The extra mass and high charge density that distinguish the neurofilament
 charged scaffolding structure suitable for interaction with other neuronal componen
 C:Comment: The boundaries of the domains between residues 70-92 and 399-402 are not
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:1-70/Domain: head <HED>
 F:92-123/Domain: coil 1a, alpha-helical rod <RIA>
 F:137-232/Domain: coil 1b, alpha-helical rod <1b>
 F:255-399/Domain: coil 2, alpha-helical rod <2>
 F:402-548/Domain: tail <TAI>
 F:402-442/Region: tail subdomain a
 F:443-548/Region: tail subdomain b

Query Match 100.0%; Score 36; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
 |||||
 Db 221 LKRVHEE 227

RESULT 12

JM0094
 neurofilament protein-L - bovine
 N:Alternate names: NF-L
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 07-May-1999
 C:Accession: JM0094
 R:Hashimoto, R.; Nakamura, Y.; Goto, H.; Wada, Y.; Sakoda, S.; Kalbuch, K.; Inagaki, M.
 Biochem. Biophys. Res. Commun. 245: 407-411, 1998
 A:Title: Domain- and site-specific phosphorylation of bovine NF-L by Rho-associated kinase
 A:Reference number: JM0094; MUID:98238650; PMID:9571164
 A:Accession: JM0094
 A:Molecule type: protein
 A:Residues: 1-554 <HMS>
 C:Comment: Domain- and site-specific phosphorylation by Rho-kinase regulates the assembly
 C:Superfamily: cytoskeletal keratin

Query Match

Best Local Similarity 100.0%; Score 36; DB 2; Length 554;
 Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
 |||||

DB 222 LKRVHEE 228

RESULT 13

B46024
 neurofilament-L subunit - quail
 C:Species: Coturnix coturnix (quail)
 C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
 C:Accession: B46024
 R:Ohara, O.; Gahara, Y.; Miyake, T.; Terakoka, H.; Kitamura, T.
 J. Cell Biol. 121: 387-395, 1993
 A:Title: Neurofilament deficiency in quail caused by nonsense mutation in neurofilament-L
 A:Reference number: A46024; MUID:93224534; PMID:8468353
 A:Accession: B46024
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-556 <OHAS>
 A:Experimental source: subsp. japonica, TKP
 A>Note: sequence inconsistent with the nucleotide translation
 A>Note: sequence extracted from NCBI backbone (NCBIN:129455, NCBIPI:129482)
 C:Superfamily: cytoskeletal keratin

Query Match

Best Local Similarity 100.0%; Score 36; DB 2; Length 556;
 Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
 |||||

DB 224 LKRVHEE 230

RESULT 14

T19995
 hypothetical protein C47D12.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T19995
 R:Gajadaty, S.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z19209
 A:Accession: T19995
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-410 <WTL>
 A:Cross-references: EMBL:Z69902; PIDD:CAA93763.1; GSPDB:GN00020; CESP:C47D12.5
 C:Experimental source: clone C47D12
 C:Genetics:
 A:Gene: CESP:C47D12.5
 A:Map position: 2

A:Introns: 32/2; 136/3; 171/3; 230/1; 331/3

Query Match

Best Local Similarity 97.2%; Score 35; DB 2; Length 410;
 Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
 |||||

DB 217 LKRVHEE 223

RESULT 15

A43554
 desmin - African clawed frog
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 13-Aug-1999
 C:Accession: A43554
 R:Herrmann, H.; Fouquet, B.; Franke, W.W.
 Development 105: 299-307, 1989
 A:Title: Expression of intermediate filament proteins during development of xenopus 1
 A:Reference number: A43554; MUID:90032404; PMID:2806128
 A:Accession: A43554
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-458 <HER>
 A:Cross-references: GB:X16842; NID:964652; PIDD:CAA34740.1; PID:964653
 C:Superfamily: cytoskeletal keratin

Query Match

Best Local Similarity 97.2%; Score 35; DB 2; Length 458;
 Pred. No. 22;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
 |||||

DB 227 LKRVHEE 233

Search completed: November 13, 2002, 13:23:25
 Job time: 12.0152 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 5.80851 seconds

(without alignments)
49,984 Million cell updates/sec

Title: US-09-856-086-2

Sequence: 1 LKKVHEE 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SWISSProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	463	1	DESM_CHICK
2	36	100.0	468	1	DESM_MESAU
3	36	100.0	468	1	DESM_MOUSE
4	36	100.0	468	1	DESM_RAT
5	36	100.0	469	1	DESM_BOVIN
6	36	100.0	469	1	DESM_HUMAN
7	36	100.0	470	1	DESM_PIG
8	36	100.0	541	1	NFL_RAT
9	36	100.0	542	1	NFL_MOUSE
10	36	100.0	543	1	NFL_HUMAN
11	36	100.0	544	1	NFL_XENLA
12	36	100.0	544	1	NFL_PIG
13	36	100.0	554	1	NFL_BOVIN
14	36	100.0	555	1	NFL_COTJA
15	35	97.2	458	1	DESM_XENLA
16	33	91.7	359	1	GFAP_CARAU
17	33	91.7	456	1	PERI_XENLA
18	33	91.7	458	1	IF3T_TORCA
19	33	91.7	465	1	VIME_HUMAN
20	33	91.7	468	1	PERI_RAT
21	33	91.7	470	1	XNIF_XENLA
22	33	91.7	471	1	PERI_HUMAN
23	33	91.7	475	1	PERI_MOUSE
24	33	91.7	475	1	DING_BACSU
25	33	91.7	3135	1	S230_PLAFO
26	32	88.9	189	1	CME2_BACSU
27	32	88.9	428	1	GFAP_BOVIN
28	32	88.9	430	1	GFAP_RAT
29	32	88.9	432	1	GFAP_HUMAN
30	31	86.1	339	1	E2B1_METJA
31	31	86.1	589	1	GLMS_METHA
32	31	86.1	766	1	STB6_YEAST
33	31	86.1	914	1	IF42_YEAST

34	31	86.1	928	1	RR44_HUMAN	09Y211 homo sapien
35	31	86.1	956	1	SVL_AQUAE	066651 aequifex aeo
36	30	83.3	197	1	SGUA_RICPR	09ZC07 ticketctsla
37	30	83.3	243	1	VIME2_CARAU	P48672 carassius a
38	30	83.3	274	1	YA03_MYCTU	083940 treponema p
39	30	83.3	285	1	MURG_LISIN	005588 mycobacteri
40	30	83.3	363	1	SEPE_MOUSE	092992 listeria in
41	30	83.3	434	1	VIME_CRRGR	09R114 mus musculu
42	30	83.3	448	1	VIME_CARAU	P48670 citicellus
43	30	83.3	450	1	PLST_CARAU	P48673 carassius a
44	30	83.3	453	1	VIME_CYPCA	P31393 carassius a
45	30	83.3	455	1	VIME_CYPCA	092155 cyprinus ca

ALIGNMENTS

RESULT 1
DESM_CHICK STANDARD; PRT; 463 AA.
ID P02542;
AC 21-JUL-1986 (Rel. 01, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desmin.
GN DES.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
ON NCBI_TaxID=9031;
OX [1]
XP TISSUE=clzard;
RC MEDLINE=84207925; PubMed=6202512;
RA Geisler N., Weber K.;
RT "The amino acid sequence of chicken muscle desmin provides a common structural model for intermediate filament proteins.";
RL EMBO J. 1:1649-1656(1982).
[2]
RN SEQUENCE OF 1-88 AND 254-415.
RX MEDLINE=83025086; PubMed=6889923;
RA Geisler N., Kaufmann E., Weber K.;
RT "Biochemical characterization of three structurally distinct domains along the protofilament unit of desmin 10 nm filaments.";
RL Cell 30:277-286(1982).
[3]
RN SEQUENCE OF 324-463.
RX MEDLINE=82037793; PubMed=6945574;
RA Geisler N., Weber K.;
RT "Comparison of the proteins of two immunologically distinct intermediate-sized filaments by amino acid sequence analysis: desmin and vimentin.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:4120-4123(1981).
[4]
RN SEQUENCE OF 366-463 FROM N.A.
RX MEDLINE=85063701; PubMed=6594672;
RA Capetanaki Y.G., Ngai J., Lazarides E.;
RT "Characterization and regulation of the expression of a gene coding for the intermediate filament protein desmin.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:6909-6913(1984).
-1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN MUSCLE CELLS. IN ADULT STRIPED MUSCLE THEY FORM A FIBROUS NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
-1- SUBUNIT: HOMOPOLYMER.
-1- SUBCELLULAR LOCATION: Cytoplasmic.
-1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation CC at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
or send an email to license@lsb-sib.ch).

DR EMBL: K02445; AAA48751.1; -.
DR PIR: A02957; DMCH.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
KM Intermediate filament; Coiled coil; Muscle protein.
FT MOD_RES 1 1 BLOCKED.
FT DOMAIN 1 99 HEAD.
FT 403 403 ROD.
FT DOMAIN 100 463 TAIL.
FT DOMAIN 100 132 COIL 1A.
FT DOMAIN 133 142 LINKER 1.
FT DOMAIN 143 243 COIL 1B.
FT 259 259 LINKER 12.
FT DOMAIN 260 278 COIL 2A.
FT DOMAIN 279 286 LINKER 2.
FT DOMAIN 287 403 COIL 2B.
FT SITE 345 345 STUTTER.
SQ SEQUENCE 463 AA; 53279 MW; F9AFC2CF9CD111C CRC64;

Query Match 100.0%; Score 36; DB 1; Length 463;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
Db 230 LKRVHEE 236
|||||||
RESULT 2
ID DESM_MESAU STANDARD; PRT; 468 AA.
AC P02541;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desmin.
GN DES.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus
NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86079506; PubMed=3855248;
RA Quax W.J., van den Broek L., Egberts W.V., Ramekiers F.,
Bloemendal H.;
RT "Characterization of the hamster desmin gene: expression and
formation of desmin filaments in nonmuscle cells after gene
transfer.";
RL Cell 43:327-338(1985).
RN [2]
RP SEQUENCE OF 171-468 FROM N.A.
RX MEDLINE=85014890; PubMed=6091127;
RA Quax W.J., van den Heuvel R., Egberts W.V., Quax-Jeukens Y.E.F.M.,
Bloemendal H.;
RT "Intermediate filament cDNAs from BHK-21 cells: demonstration of
distinct genes for desmin and vimentin in all vertebrate classes.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5970-5974(1984).
CC -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
or send an email to license@lsb-sib.ch).

DR EMBL: K02407; AAA37071.1; -.
DR EMBL: M12104; AAA37072.1; -.
DR EMBL: M12102; AAA37072.1; JOINED.
DR EMBL: M12103; AAA37072.1; JOINED.
DR PIR: A02956; DMHY.
DR PIR: A24783; A24783.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
KM Intermediate filament; Coiled coil; Muscle protein.
FT INIT_MET 0 0
FT DOMAIN 1 106 HEAD.
FT 107 410 ROD.
FT DOMAIN 411 468 TAIL.
FT DOMAIN 107 139 COIL 1A.
FT DOMAIN 140 149 LINKER 1.
FT DOMAIN 150 250 COIL 1B.
FT DOMAIN 251 266 LINKER 12.
FT DOMAIN 267 285 COIL 2A.
FT 286 293 LINKER 2.
FT DOMAIN 294 410 COIL 2B.
FT SITE 352 352 STUTTER.
SQ SEQUENCE 468 AA; 53307 MW; 39159431C5908B9 CRC64;

Query Match 100.0%; Score 36; DB 1; Length 468;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
Db 237 LKRVHEE 243
|||||||
RESULT 3
ID DESM_MOUSE STANDARD; PRT; 468 AA.
AC P31001;
DT 01-JUL-1993 (Rel. 26, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desmin.
GN DES.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94165148; PubMed=8120103;
RA Li H., Choudhary S.K., Milner D.J., Munir M.I., Kulsk I.R.,
Capetanaki Y.;
RT "Inhibition of desmin expression blocks myoblast fusion and
interferes with the myogenic regulators MyoD and myogenin.";
RL J. Cell Biol. 124:827-841(1994).
RN [2]
RP SEQUENCE OF 1-39 FROM N.A.
RX STRAIN-BALB/c; TISSUE=Spleen;
MEDLINE=93181210; PubMed=8382796;
RA Li H., Capetanaki Y.;
RT "Regulation of the mouse desmin gene: transactivated by MyoD,
Nucleic acids Res. 21:335-343(1993).
CC -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
CC -1- SUBUNIT: HOMOPOLYMER.

```

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L23550; -; NOT_ANNOTATED_CDS.
DR EMBL: Z18892; CAA79330.1; -.
DR PIR: S31404; S31404.
DR SWISS-2DPAGE: P31001; MOUSE.
DR MGD: MGI:94885; Des.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; Filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Muscle protein.
FT INIT MET 0 0 BY SIMILARITY.
FT DOMAIN 1 107 HEAD.
FT DOMAIN 108 410 ROD.
FT DOMAIN 411 468 TAIL.
FT DOMAIN 108 139 COIL_1A.
FT DOMAIN 140 149 LINKER 1.
FT DOMAIN 150 250 COIL_1B.
FT DOMAIN 251 266 LINKER 12.
FT DOMAIN 267 285 COIL_2A.
FT DOMAIN 286 293 LINKER 2.
FT DOMAIN 294 410 COIL_2B.
FT SITE 352 352 STUTTER.
FT SITE 44 47 POLY-SER.
SQ SEQUENCE 468 AA; 53366 MW; FD276059E7E3BFB CRC64;

Query Match 100.0%; Score 36; DB 1; Length 468;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 237 LKRVHEE 243

RESULT 4
DESM_BOVIN STANDARD; PRT; 468 AA.
ID DESM_BOVIN
AC P48675;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desmin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holtzman; TISSUE=Aorta;
RX MEDLINE=94114566; PubMed=8286410;
RA van Groningen J.J.M., Bloemers H.P.J., Swart G.W.M.;
RT "Rat desmin gene structure and expression.";
RL Blochin, Biophys. Acta 1217:107-109(1994).
CC -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
CC MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
CC NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
CC MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
CC -1- SUBUNIT: HOMOPOLYMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L23550; -; NOT_ANNOTATED_CDS.
DR EMBL: Z18892; CAA79330.1; -.
DR PIR: S31404; S31404.
DR SWISS-2DPAGE: P31001; MOUSE.
DR MGD: MGI:94885; Des.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; Filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Muscle protein.
FT INIT MET 0 0 BY SIMILARITY.
FT DOMAIN 1 107 HEAD.
FT DOMAIN 108 410 ROD.
FT DOMAIN 411 468 TAIL.
FT DOMAIN 108 139 COIL_1A.
FT DOMAIN 140 149 LINKER 1.
FT DOMAIN 150 250 COIL_1B.
FT DOMAIN 251 266 LINKER 12.
FT DOMAIN 267 285 COIL_2A.
FT DOMAIN 286 293 LINKER 2.
FT DOMAIN 294 410 COIL_2B.
SQ SEQUENCE 468 AA; 53366 MW; FD276059E7E3BFB CRC64;

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X73524; CAA51920.1; -.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; Filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Muscle protein.
FT INIT MET 0 0 BY SIMILARITY.
FT DOMAIN 1 107 HEAD.
FT DOMAIN 108 410 ROD.
FT DOMAIN 411 468 TAIL.
FT DOMAIN 108 139 COIL_1A.
FT DOMAIN 140 149 LINKER 1.
FT DOMAIN 150 250 COIL_1B.
FT DOMAIN 251 266 LINKER 12.
FT DOMAIN 267 285 COIL_2A.
FT DOMAIN 286 293 LINKER 2.
FT DOMAIN 294 410 COIL_2B.
SQ SEQUENCE 468 AA; 53325 MW; EBD6675A5E7FA8B CRC64;

Query Match 100.0%; Score 36; DB 1; Length 468;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 237 LKRVHEE 243

RESULT 5
DESM_BOVIN STANDARD; PRT; 469 AA.
ID DESM_BOVIN
AC O62654; O62655;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desmin.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Muscle;
RA Chikuni K., Tanabe R., Muroya S.;
RT "Desmin structure as related to meat tenderness.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
CC MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
CC NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
CC MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
CC -1- SUBUNIT: HOMOPOLYMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB011675; BAA25135.1; -.
DR EMBL: AB011673; BAA25133.1; -.
DR InterPro: IPR001664; IF.

```

DR Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF: 1.
 KW Intermediate filament; Coiled coil; Muscle protein.
 FT INT MET 0 0 BY SIMILARITY.
 FT DOMAIN 1 107 HEAD.
 FT DOMAIN 108 411 ROD.
 FT DOMAIN 412 469 TAIL.
 FT DOMAIN 108 140 COIL 1A.
 FT DOMAIN 141 150 LINKER 1.
 FT DOMAIN 151 251 COIL 1B.
 FT DOMAIN 252 267 LINKER 12.
 FT DOMAIN 268 286 COIL 2A.
 FT DOMAIN 287 294 LINKER 2.
 FT DOMAIN 295 411 COIL 2B.
 FT SITE 353 353 STUTTER.
 FT DOMAIN 44 47 POLY-SER.
 SO SEQUENCE 469 AA: 53400 MW: C7275DC3E528DFF5 CRC64;
 Query Match 100.0%; Score 36; DB 1; Length 469;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 LKKVHEE 7
 Db 238 LKKVHEE 244
 RESULT 6
 DESM_HUMAN
 ID DESM_HUMAN STANDARD: PRT: 469 AA.
 AC P17661; Q13787;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Desmin.
 GN DES.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A. PubMed-2673923;
 RX MEDLINE-89378751; PubMed-2673923;
 RA Li Z., Lilienbaum A., Butler-Browne G., Paulin D.;
 RT "Human desmin-coding gene: complete nucleotide sequence,
 RT characterization and regulation of expression during myogenesis and
 RT development";
 RL Gene 78:243-254(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Li Z., Paulin D.;
 RL Submitted (XXX-1991) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RX MEDLINE-96384956; PubMed-8792816;
 RA Viscart P., Dupret J.M., Hazan J., Li Z., Gyapay G., Krishnamoorthy R.,
 RA Weisenbach J., Fardieu M., Paulin D.;
 RT "Human desmin gene: cDNA sequence, regional localization and
 RT exclusion of the locus in a familial desmin-related myopathy";
 RL Hum. Genet. 98:422-429(1996).
 RN [4]
 RP SEQUENCE FROM N.A. AND VARIANTS CSM PRO-336; PRO-359 AND ILE-392.
 RX MEDLINE-98361171; PubMed-9697706;
 RA Goldfarb L.G., Park K.-Y., Cervenakova L., Gorokhova S., Lee H.-S.,
 RA Vasconcelos O., Nagle J.W., Semino-Mora C., Sivakumar K.,
 RA Dalakas M.C.;
 RT "Missense mutations in desmin associated with familial cardiac and
 RT skeletal myopathy";
 RL Nat. Genet. 19:402-403(1998).
 RP VARIANT CSM 172-ARG--GLU-178 DEL.
 RN [5]
 RX MEDLINE-98409654; PubMed-9736733;

RA Munoz-Marmol A.M., Strasser G., Isamat M., Coulombe P.A., Yang Y.,
 RA Roca X., Vela E., Mate J.L., Coll J., Fernandez-Figueras M.T.,
 RA Navas-Palacios J.J., Ariza A., Fuchs E.;
 RT "A dysfunctional desmin mutation in a patient with severe generalized
 RT myopathy";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:11312-11317(1998).
 RN [6]
 RP VARIANT HDM PRO-344.
 RX MEDLINE-20014709; PubMed-10545598;
 RA Sjoeborg G., Saavedra-Matiz C.A., Rosen D.R., Wajsmann E.M., Borg K.,
 RA Horowitz S.H., Sejersen T.;
 RT "A missense mutation in the desmin rod domain is associated with
 RT autosomal dominant distal myopathy, and exerts a dominant negative
 RT effect on filament formation";
 RL Hum. Mol. Genet. 8:2191-2198(1999).
 CC -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
 CC MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
 CC NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
 CC MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
 CC -1- SUBUNIT: HOMOPOLYMER.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- DISEASE: DEFECTS IN DES ARE THE CAUSE OF DESMIN-RELATED MYOPATHY,
 CC A FAMILIAL CARDIAC AND SKELETAL MYOPATHY (CSM). CSM IS
 CC CHARACTERIZED BY SKELETAL MUSCLE WEAKNESS ASSOCIATED WITH CARDIAC
 CC CONDUCTION BLOCKS, ARRHYTHMIAS AND RESTRICTIVE HEART FAILURE, AND
 CC BY INTRACYTOLASMIC ACCUMULATION OF DESMIN-REACTIVE DEPOSITS IN
 CC CARDIAC AND SKELETAL MUSCLE CELLS. A DESMIN-RELATED MYOPATHY CAN
 CC HAVE A DISTAL ONSET, IT IS THEN KNOWN AS HEREDITARY DISTAL
 CC MYOPATHY (HDM).
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: M63391; AAA99221.1; -;
 CC DR EMBL: U59167; AAC50680.1; -;
 CC DR PIR: J60063; DMHD.
 CC DR HSC-2DPAGE: P17661; HUMAN.
 CC DR Gene: HGNC:2770; DES.
 CC DR MIM: 601419; -;
 CC DR MIM: 601419; -;
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; filament; 1.
 DR PROSITE: PS00226; IF: 1.
 KW Intermediate filament; Coiled coil; Muscle protein; Disease mutation.
 FT INT MET 0 0
 FT DOMAIN 1 107 HEAD.
 FT DOMAIN 108 411 ROD.
 FT DOMAIN 412 469 TAIL.
 FT DOMAIN 108 140 COIL 1A.
 FT DOMAIN 141 150 LINKER 1.
 FT DOMAIN 151 251 COIL 1B.
 FT DOMAIN 252 267 LINKER 12.
 FT DOMAIN 268 286 COIL 2A.
 FT DOMAIN 287 294 LINKER 2.
 FT DOMAIN 295 411 COIL 2B.
 FT DOMAIN 44 47 MISSING (IN CSM; SEVERE FORM).
 FT VARIANT /FtId-VAR_009188.
 FT VARIANT /FtId-VAR_007900.
 FT VARIANT /FtId-VAR_007900.
 FT VARIANT /FtId-VAR_009189.
 FT VARIANT A -> P (IN CSM; HETEROZYGOUS WITH ILE-391
 FT GIVES A SEVERE CSM CHILDHOOD-ONSET).
 FT /FtId-VAR_007901.
 FT N -> I (IN CSM; HETEROZYGOUS WITH PRO-358
 FT GIVES A SEVERE CSM CHILDHOOD-ONSET).
 FT /FtId-VAR_007902.

```

FT CONFLICT 22 24 GFF -> VFS (IN REF. 1 AND 2).
FT CONFLICT 38 38 G -> P (IN REF. 1 AND 2).
FT CONFLICT 118 122 FANYI -> SPIYM (IN REF. 1 AND 2).
FT CONFLICT 134 134 MISSING (IN REF. 1, 2 AND 3).
SQ SEQUENCE 469 AA; 53404 MW; 6A38116859A091B8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 36; DB 1; Length 469;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
DB 238 LKRVHEE 244

RESULT 7
DESM_PIG STANDARD; PRT; 470 AA.
AC P02540; 062656;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Desmin.
GN DES.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9623;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Chikuni K., Tanabe R., Muroya S.;
RT "Desmin structure as related to meat tenderness.";
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Tissue=Longissimus muscle;
RA Beunen N.D., Hall A.D., Gallagher A., Chang K.-C.;
RT "A polymorphic C-repeat at the porcine desmin locus with an effect on meat quality.";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Longissimus muscle;
RA Beunen N.D., Hall A.D., Gallagher A., Chang K.-C.;
RT "A polymorphic C-repeat at the porcine desmin locus with an effect on meat quality.";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 333-470.
RX MEDLINE=82037793; PubMed=6945574;
RA Geisler N., Weber K.;
RT "Comparison of the proteins of two immunologically distinct intermediate-sized filaments by amino acid sequence analysis: desmin and vimentin.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:4120-4123(1981).
CC -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB011676; BAA25136.1; -
DR EMBL; AB011674; BAA25134.1; -
DR EMBL; AF136188; AAD46492.1; -

```

```

DR EMBL; AF363284; AAK51087.1; -
DR PIR; A02955; DMFG.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Muscle protein.
FT INT_MET 0
FT DOMAIN 1 108
FT DOMAIN 109 412
FT DOMAIN 413 470
FT DOMAIN 109 141
FT DOMAIN 142 151
FT DOMAIN 152 252
FT DOMAIN 253 268
FT DOMAIN 269 287
FT DOMAIN 288 295
FT DOMAIN 296 412
FT SITE 354 354
FT SITE 354 354
FT DOMAIN 44 47
SQ SEQUENCE 470 AA; 53497 MW; A2ABF7A8DB65DD12 CRC64;

Query Match
Best Local Similarity 100.0%; Score 36; DB 1; Length 470;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
DB 239 LKRVHEE 245

RESULT 8
NFL_RAT STANDARD; PRT; 541 AA.
AC P19527; 063367;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NFL-L).
GN NFL OR NFL OR NF68.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90184052; PubMed=2516804;
RA Chin S.S., Liem R.K.H.;
RT "Expression of rat neurofilament proteins NF-L and NF-M in transfected non-neuronal cells.";
RL Eur. J. Cell Biol. 50:475-490(1989).
RN [2]
RP SEQUENCE OF 197-483 FROM N.A.
RX MEDLINE=85252830; PubMed=3925999;
RA Julien J.-P., Ramachandran K., Grosveid F.;
RT "Cloning of a cDNA encoding the smallest neurofilament protein from the rat.";
RL Biochim. Biophys. Acta 825:398-404(1985).
RN [3]
RP SEQUENCE OF 1-10 FROM N.A.
RX MEDLINE=95264348; PubMed=7745611;
RA Reeben M., Neuman T., Palgi J., Palm K., Paalme V., Saarma M.;
RT "Characterization of the rat light neurofilament (NF-L) gene promoter and identification of NF and cAMP responsive regions.";
RL J. Neurosci. Res. 40:177-188(1995).
RN [4]
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=93346421; PubMed=8344946;
RA Dong D.L.-Y., Xu Z.-S., Cheyrier M.R., Cotter R.J., Cleveland D.W., Hart G.W.;
RT "Glycosylation of mammalian neurofilaments. Localization of multiple O-linked N-acetylglucosamine moieties on neurofilament polypeptides L and M.";

```

RL J. Biol. Chem. 268:16679-16687(1993).

CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- PTM: O-GLYCOSYLATED; CONTAINS THREE N-ACETYLGLUCOSAMINE SIDE
CC CHAINS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb-sib.ch).
CC -----
CC EMBL: AF031880; AAB87069.1; -
CC EMBL: M25638; AAA41694.1; -
CC EMBL: X53981; CA37931.1; -
CC PIR: A21762; A21762.
CC GLCOSuitedB: P19527; -
CC InterPro: IPR001664; IF.
CC Pfam: PF00038; filament.1.
CC PROSITE: PS00226; IF.1.
CC KW Intermediate filament; Coiled coil; Neurone; Glycoprotein.
CC FT INT_MET 0 0
CC FT DOMAIN 1 92 HEAD.
CC FT DOMAIN 93 396 ROD.
CC FT DOMAIN 397 541 TAIL.
CC FT DOMAIN 93 124 COIL 1A.
CC FT DOMAIN 125 137 LINKER 1.
CC FT DOMAIN 138 233 COIL 1B.
CC FT DOMAIN 234 252 LINKER 12.
CC FT DOMAIN 253 271 COIL 2A.
CC FT DOMAIN 272 280 LINKER 2.
CC FT DOMAIN 281 396 COIL 2B.
CC FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
CC FT DOMAIN 444 541 TAIL, SUBDOMAIN B (ACIDIC).
CC FT CARBOHYD 20 20
CC FT CARBOHYD 26 26
CC FT SITE 381 391
CC FT CONFLICT 197 202
CC FT CONFLICT 399 399
CC FT CONFLICT 476 476
CC FT CONFLICT 480 483
CC FT SEQUENCE 541 AA; 61204 MW; 0D17839AF226918A CnC64;
CC
CC Query Match 100.0%; Score 36; DB 1; Length 541;
CC Best Local Similarity 100.0%; Pred. No. 5.9; Indels 0; Gaps 0;
CC Matches 7; Conservative 0; Mismatches 0;
CC
CC Oy 1 LKRVHEE 7
CC Db 222 LKRVHEE 228
CC
CC RESULT 9
CC ID_NFL_MOUSE STANDARD; PRT; 542 AA.
CC AC P08551;
CC DT 01-AUG-1988 (Rel. 08, Created)
CC DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Neurofilament triplet L protein (68 kDa neurofilament protein)

DE (Neurofilament light polypeptide) (NF-L).

GN NEFL OR NEFL.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-87064433; PubMed-3785173;
RA Lewis S.A., Cowan N.J.;
RT "Anomalous placement of introns in a member of the intermediate
RT filament multigene family: an evolutionary conundrum.";
RL Mol. Cell. Biol. 6:1529-1534(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-87158637; PubMed-3103856;
RA Julien J.-P., Meyer D., Flavell D., Hurst J., Grosfeld F.;
RT "Cloning and developmental expression of the murine neurofilament
RT gene family.";
RL Brain Res. 387:243-250(1986).
RN [3]
RP SEQUENCE OF 241-542 FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-85131334; PubMed-3919033;
RA Lewis S.A., Cowan N.J.;
RT "Genetics, evolution, and expression of the 68,000-mol-wt
RT neurofilament protein: isolation of a cloned cDNA probe.";
RL J. Cell Biol. 100:843-850(1985).
RN [4]
RP SEQUENCE OF 1-27 FROM N.A.
RX MEDLINE-91060592; PubMed-2246261;
RA Nakhshibi K., Ikenaka K., Wada K., Tamura T.A., Furutachi T.,
RA Mikoshiba K.;
RT "Structure of the 68-kDa neurofilament gene and regulation of its
RT expression.";
RN J. Biol. Chem. 265:19786-19791(1990).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb-sib.ch).
CC -----
CC EMBL: X02165; CAB51616.1; -
CC EMBL: M20480; AAA39814.1; -
CC EMBL: M13016; AAA39810.1; -
CC EMBL: M55423; AAA39812.1; -
CC PIR: A25227; QFMSL.
CC WGI: 97313; NFL.
CC InterPro: IPR001664; IF.
CC Pfam: PF00038; filament.1.
CC PROSITE: PS00226; IF.1.
CC KW Intermediate filament; Coiled coil; Neurone; Glycoprotein.
CC FT INT_MET 0 0
CC FT DOMAIN 1 92 HEAD.

```

FT DOMAIN 93 396 ROD.
FT DOMAIN 397 542 TAIL.
FT DOMAIN 93 124 COIL 1A.
FT DOMAIN 125 137 LINKER 1.
FT DOMAIN 138 233 COIL 1B.
FT DOMAIN 234 232 LINKER 12.
FT DOMAIN 253 271 COIL 2A.
FT DOMAIN 272 280 LINKER 2.
FT DOMAIN 281 396 COIL 2B.
FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
FT DOMAIN 444 542 TAIL, SUBDOMAIN B (ACIDIC).
FT CARBOHYD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
FT CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
FT SITE 391 391 EPIPODE (RECOGNIZED BY IF-SPECIFIC MONOCLONAL ANTIBODY).
FT CONFLICT 5 5 Y -> S (IN REF. 1).
FT CONFLICT 6 6 Y -> I (IN REF. 1).
FT CONFLICT 64 64 M -> K (IN REF. 1).
FT CONFLICT 72 72 V -> L (IN REF. 2).
FT CONFLICT 98 98 D -> H (IN REF. 1).
FT CONFLICT 194 194 R -> A (IN REF. 1).
FT CONFLICT 202 202 MISSING (IN REF. 2).
FT CONFLICT 239 239 Y -> I (IN REF. 1).
SQ SEQUENCE 542 AA; 61448 MW; 8EE9B8C6F0831D8C CRC64;

Query Match 100.0%; Score 36; DB 1; Length 542;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LKRVHEE 7
Db 222 LKRVHEE 228
|||||
NF-L_HUMAN STANDARD; PRT; 543 AA.
AC P07196; Q16154;
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
DE NEFL OR NEFL OR NF68.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA JULIEN J.-P., GROSVELD F., YAZDANBASH K., FLAVELL D., MEIJER D.,
RA MUSHYANSKI W.;
RA "The structure of a human neurofilament gene (NF-L): a unique exon-
intron organization in the intermediate filament gene family.";
RL Blochm. Biophys. Acta 909:10-20(1987).
RN [2]
RP SEQUENCE OF 1-10 FROM N.A.
RA MEDLINE=94235564; PubMed=8180132;
RA "Ap-1 and Krox-24 transcription factors activate the neurofilament
light gene promoter in p19 embryonal carcinoma cells.";
RL Cell Growth Differ. 5:187-196(1994).
RN [3]
RP VARIANT CMT2E PRO-332.
RA MEDLINE=20307176; PubMed=10841809;
RA MERSIYANOVA I.V., PEREPELOVA A.V., POLYAKOV A.V., SITNIKOV V.F.,
RA DADALI E.L., OPATIN R.B., PETLIN A.N., EYSGRAFOV O.V.;
RA "A new variant of Charcot-Marie-Tooth disease type 2 is probably the
result of a mutation in the neurofilament-light gene.";
RL Am. J. Hum. Genet. 67:37-46(2000).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.

```

```

CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- DISEASE: DEFECTS IN NEFL ARE A CAUSE OF CHARCOT-MARIE-TOOTH
CC DISEASE TYPE 2E (CMT2E).
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -1- DATABASE: NAME-INHERITED peripheral neuropathies mutation db;
CC WWW="http://molgen-www.uia.ac.be/CMTmutations/".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05608; CAA29097.1; -
DR EMBL; S70369; AAD14057.1; -
DR PIR; S07144; S07144.
DR Gene; HGNC:7739; NEFL.
DR MIM; 162280; -.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neurone; Glycoprotein;
KW Disease mutation; Charcot-Marie-Tooth disease.
FT INTR_MET 0 0
FT DOMAIN 1 91 HEAD.
FT DOMAIN 92 396 ROD.
FT DOMAIN 397 543 TAIL.
FT DOMAIN 92 123 COIL 1A.
FT DOMAIN 124 136 LINKER 1.
FT DOMAIN 137 234 COIL 1B.
FT DOMAIN 235 252 LINKER 12.
FT DOMAIN 253 271 COIL 2A.
FT DOMAIN 272 280 LINKER 2.
FT DOMAIN 281 396 COIL 2B.
FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
FT DOMAIN 444 543 TAIL, SUBDOMAIN B (ACIDIC).
FT CARBOHYD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
FT CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
FT SITE 391 391 EPIPODE (RECOGNIZED BY IF-SPECIFIC MONOCLONAL ANTIBODY).
FT VARIANT 332 332 Q -> P (IN CMT2E).
FT FTID=VAR_009703.
SQ SEQUENCE 543 AA; 61645 MW; 7A0F1ADD5BD22F6 CRC64;

Query Match 100.0%; Score 36; DB 1; Length 543;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LKRVHEE 7
Db 223 LKRVHEE 229
|||||
NF-L_XENLA STANDARD; PRT; 544 AA.
AC P35616;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (Neurofilament light polypeptide) (NF-
L).
OS Xenopus laevis (African clawed frog).

```


CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 CC Xenopodidae; Xenopus.
 CC NCBI_TaxID=8353;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92356194; PubMed=1494944;
 RA Charnas L.R., Szaro B.G., Garner H.;
 RT "Identification and developmental expression of a novel low molecular
 RT weight neuronal intermediate filament protein expressed in Xenopus
 RT laevis.";
 RL J. Neurosci. 12:3010-3024(1992).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M86654; AAA83018.1; -
 DR PIR; B44841; B44841.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; Filament; 1.
 DR PROSITE; PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neurone.
 FT DOMAIN 1 87 HEAD.
 FT DOMAIN 88 390 ROD.
 FT DOMAIN 391 544 TAIL.
 FT DOMAIN 88 119 COIL 1A.
 FT DOMAIN 120 132 COIL 1B.
 FT DOMAIN 133 228 COIL 1B.
 FT DOMAIN 229 246 COIL 1B.
 FT DOMAIN 247 265 COIL 2A.
 FT DOMAIN 266 274 COIL 2A.
 FT DOMAIN 275 390 COIL 2B.
 FT DOMAIN 391 435 TAIL, SUBDOMAIN A.
 FT DOMAIN 436 544 TAIL, SUBDOMAIN B (ACIDIC).
 FT DOMAIN 441 538 GLU-RICH.
 FT DOMAIN 464 469 POLY-GLU.
 SQ SEQUENCE 544 AA; 61861 MW; 76D911B896E97201 CRC64;
 Query Match 100.0%; Score 36; DB 1; Length 544;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LKRVHEE 7
 Db 217 LKRVHEE 223
 RESULT 12
 NFL_PIG STANDARD; PRT; 548 AA.
 AC P02347;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L).
 GN NEFL.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85154583; PubMed=3920075;
 RA Geisler N., Plessmann U., Weber K.;
 RT "The complete amino acid sequence of the major mammalian
 RT neurofilament protein (NF-L)."
 RL FEBS Lett. 182:475-478(1985).
 RN [2]
 RP SEQUENCE OF 1-82 AND 278-548.
 RA Geisler N., Kaufmann E., Fischer S., Plessmann U., Weber K.;
 RT "Neurofilament architecture combines structural principles of
 RT intermediate filaments with carboxy-terminal extensions increasing
 RT in size between triplet proteins.";
 RL EMBL J. 2:1295-1302(1983).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC
 CC PIR; A02963; OEPGL.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; Filament; 1.
 DR PROSITE; PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neurone; Glycoprotein.
 FT DOMAIN 1 91 HEAD.
 FT DOMAIN 92 395 ROD.
 FT DOMAIN 396 548 TAIL.
 FT DOMAIN 92 123 COIL 1A.
 FT DOMAIN 124 136 COIL 1B.
 FT DOMAIN 137 232 COIL 1B.
 FT DOMAIN 233 251 COIL 1B.
 FT DOMAIN 252 270 COIL 2A.
 FT DOMAIN 271 279 COIL 2A.
 FT DOMAIN 280 395 COIL 2B.
 FT DOMAIN 396 442 TAIL, SUBDOMAIN A.
 FT DOMAIN 443 548 TAIL, SUBDOMAIN B (ACIDIC).
 FT CARBOHYD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
 FT CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
 FT SITE 380 390 EPTOPE (RECOGNIZED BY IF-SPECIFIC
 FT MONOCLONAL ANTIBODY).
 FT OR K.
 SQ SEQUENCE 322 322 83044813637AC739 CRC64;
 Query Match 100.0%; Score 36; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LKRVHEE 7
 Db 221 LKRVHEE 227
 RESULT 13
 NFL_BOVIN STANDARD; PRT; 554 AA.
 AC P02548; P79127;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 DE (Neurofilament light polypeptide) (NF-L) (Micro glutamic acid-rich
 DE protein).
 GN NEFL.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Holstein; TISSUE-Brain;
 RA Hall W.D., Zhang L., Balin B.J., Sprinkle T.J.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 468-554.
 RX MEDLINE=85154567; PubMed=3884373;
 RA Ito T., Okuyama T.;
 RT "Brain micro glutamic acid-rich protein is the C-terminal endpiece of
 the neurofilament 68-kDa protein as determined by the primary
 sequence.";
 RL FEBS Lett. 182:389-392(1985).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U83919; AAB41543.1; -
 DR PIR: A02964; CEB0.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF000038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neutrone.
 FT INIT_MET 0 0
 FT DOMAIN 1 92
 FT DOMAIN 93 396
 FT DOMAIN 397 554
 FT DOMAIN 554 554
 FT DOMAIN 93 124
 FT DOMAIN 125 137
 FT DOMAIN 138 233
 FT DOMAIN 234 252
 FT DOMAIN 253 271
 FT DOMAIN 272 280
 FT DOMAIN 281 396
 FT DOMAIN 397 443
 FT DOMAIN 444 554
 FT DOMAIN 554 554
 FT CONFLICT 494 500
 FT CONFLICT 509 509
 FT SEQUENCE 554 AA; 62514 MW; D772B81CA2C31C1A CRC64;
 SO
 Query Match 100.0%; Score 36; DB 1; Length 554;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LKRVHEE 7
 DB 223 LKRVHEE 228
 RESULT 14
 ID NCBI_COTJA
 AC 002916;
 STANDARD; PRT; 555 AA.

DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (Neurofilament light polypeptide) (NF-
 L).
 GN NEFL.
 OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Coturnix.
 OC NCBI_TaxID=93934;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=93224534; PubMed=8468353;
 RA Ohara O., Gahara Y., Miyake T., Teroka H., Kitamura T.;
 RT "Neurofilament deficiency in quail caused by nonsense mutation in
 neurofilament-L gene.";
 RL J. Cell Biol. 121:387-395(1993).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D13223; BAA02504.1; -
 DR EMBL: D13222; BAA02503.1; ALT_TERM.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF000038; filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neutrone.
 FT INIT_MET 0 0
 FT DOMAIN 1 93
 FT DOMAIN 94 397
 FT DOMAIN 398 555
 FT DOMAIN 555 555
 FT DOMAIN 94 125
 FT DOMAIN 126 138
 FT DOMAIN 139 234
 FT DOMAIN 235 253
 FT DOMAIN 254 272
 FT DOMAIN 273 281
 FT DOMAIN 282 397
 FT DOMAIN 398 444
 FT DOMAIN 445 555
 FT SEQUENCE 555 AA; 62282 MW; 9B957ABD8E8A7712 CRC64;
 SO
 Query Match 100.0%; Score 36; DB 1; Length 555;
 Best Local Similarity 100.0%; Pred. No. 6.1;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LKRVHEE 7
 DB 223 LKRVHEE 229
 RESULT 15
 ID DESM_XENIA

```

ID DESM_XENLA STANDARD: PRT: 458 AA.
AC P23239;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desmin.
GN DES.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=83355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90032404; PubMed=2806128;
RA Herrmann H., Fouquet B., Franke W.W.;
RT "Expression of intermediate filament proteins during development of
RT Xenopus laevis. II. Identification and molecular characterization of
RT desmin."
RL Development 105:299-307(1989).
CC -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN
CC MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS
CC NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA
CC MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
CC -1- SUBUNIT: HOMOPOLYMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X16842; GAA34740.1;
DR PIR: A43554; A43554.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; Filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Muscle protein.
FT DOMAIN 1 96 HEAD.
FT DOMAIN 97 400 ROD.
FT DOMAIN 401 458 TAIL.
FT DOMAIN 97 129 COIL 1A.
FT DOMAIN 130 139 LINKER 1.
FT DOMAIN 140 240 COIL 1B.
FT DOMAIN 241 256 LINKER 12.
FT DOMAIN 257 275 COIL 2A.
FT DOMAIN 276 283 LINKER 2.
FT DOMAIN 284 400 COIL 2B.
FT SITE 342 342 STUTTER.
SQ SEQUENCE 458 AA; 52802 MM; E96167E1D8FDEB74 CRC64;

```

```

Query Match 97.2%; Score 35; DB 1; Length 458;
Best Local Similarity 85.7%; Pred. No. 8;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LKRVHEE 7
| | | | |
DB 227 LKRVHEE 233

```

Search completed: November 13, 2002, 13:16:50
Job time : 6.80851 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 22.4894 Seconds
(without alignments)
64.134 Million cell updates/sec

Title: US-09-856-086-2
Perfect score: 36
Sequence: 1 LKRVHEE 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_Archea:*
2: SP_Bacteria:*
3: SP_Fungi:*
4: SP_Human:*
5: SP_Invertebrate:*
6: SP_Mammal:*
7: SP_mnc:*
8: SP_Organelle:*
9: SP_Phage:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*
15: SP_Virus:*
16: SP_Bacteria:*
17: SP_Archea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	73	4	Q9GZRG6
2	36	100.0	386	4	Q8RCR7
3	36	100.0	448	13	073665
4	36	100.0	455	13	073587
5	36	100.0	470	4	Q9UHN5
6	36	100.0	470	4	Q9H319
7	36	100.0	470	4	Q8TD99
8	36	100.0	470	4	Q8TCX2
9	36	100.0	470	4	Q8TAC4
10	35	97.2	410	5	Q18665
11	35	97.2	459	13	Q8UWF1
12	35	97.2	807	10	Q48724
13	34	94.4	60	13	Q91377
14	34	94.4	68	13	Q91355
15	34	94.4	1937	13	Q91BD4
16	33	91.7	275	4	Q96H12

17	33	91.7	275	11	Q9CY5	Q9CY5 mus musculus
18	33	91.7	275	11	Q9CR78	Q9CR78 mus musculus
19	33	91.7	410	4	Q96ML2	Q96ML2 homo sapien
20	33	91.7	450	13	Q9DDB3	Q9DDB3 scyllorhinu
21	33	91.7	459	13	Q8UYC8	Q8UYC8 dabola russ
22	33	91.7	471	13	P87360	P87360 brachydania
23	33	91.7	475	4	Q9H8F5	Q9H8F5 homo sapien
24	33	91.7	495	4	Q96CRO	Q96CRO homo sapien
25	33	91.7	541	4	Q9AHN0	Q9AHN0 homo sapien
26	33	91.7	598	2	Q9AHJ8	Q9AHJ8 lacobactill
27	33	91.7	703	4	Q96MT8	Q96MT8 homo sapien
28	33	91.7	823	16	Q9PM57	Q9PM57 campylobact
29	33	91.7	847	2	Q93U57	Q93U57 carsonella
30	33	91.7	314	5	Q25994	Q25994 plasmodium
31	33	91.7	314	5	Q9GTR4	Q9GTR4 plasmodium
32	32	88.9	144	10	Q9C651	Q9C651 arabidopsis
33	32	88.9	243	16	Q971M4	Q971M4 clostridium
34	32	88.9	317	4	Q9UFD0	Q9UFD0 homo sapien
35	32	88.9	400	16	Q25446	Q25446 heliobacte
36	32	88.9	400	16	Q9ZL93	Q9ZL93 heliobacte
37	32	88.9	421	11	Q9Z250	Q9Z250 rattus norv
38	32	88.9	430	11	Q9H103	Q9H103 rattus norv
39	32	88.9	432	4	Q96P18	Q96P18 homo sapien
40	32	88.9	473	13	Q9DDC8	Q9DDC8 lampetra fl
41	32	88.9	473	13	Q90441	Q90441 brachydania
42	32	88.9	501	11	Q9CX98	Q9CX98 mus musculus
43	32	88.9	582	17	Q93737	Q93737 pyrococcus
44	31	86.1	111	5	Q85Y83	Q85Y83 drosophila
45	31	86.1	146	16	Q8XKNS	Q8XKNS clostridium

ALIGNMENTS

RESULT 1

ID Q9GZRG6 PRELIMINARY; PRT; 73 AA.
AC Q9GZRG6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Mutant desmin CSM-7 (Mutant desmin CSM-6) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Park K.-Y., Dalakas M.C., Goebel H.H., Ferrans V.J., Semino-Mora C.,
RA Litvak S., Takeda K., Goldfarb L.G.;
RA Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF189281; AAC34459.1; -;
DR EMBL; AF189280; AAC34459.1; -;
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
FT NON_TER
FT NON_TER
SQ SEQUENCE 73 AA; 8537 MW; B062E72B45FE955 CRC64;

Query Match 100.0%; Score 36; DB 4; Length 73;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 46 LKRVHEE 52

RESULT 2
Q8RCR7 PRELIMINARY; PRT; 386 AA.
AC Q8RCR7;
DT 01-JUN-2002 (TREMBLrel. 21, Created)

```

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Hypothetical 43.8 kDa protein (Fragment).
GN DKEZ761K0922.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MYOGDIAL;
RA Wambolt R., Heubner D., Mewes H.W., Well B., Wiemann S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ13644; CAD28456.1; -
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 386 AA; 43779 MW; F293388B200C7B65 CRC64;

Query Match
Best Local Similarity 100.0%; Score 36; DB 4; Length 386;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
Db 66 LKRVHEE 72

RESULT 3
ID 073665 PRELIMINARY; PRT; 448 AA.
AC 073665;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Desmin (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=MUSCLE;
RA Chikuni K.;
RT "Partial sequence of chicken desmin.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
DR EMBL; AB011672; BAA25132.1; -
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Coiled coil; Intermediate filament.
FT NON_TER
SQ SEQUENCE 448 AA; 51664 MW; 198EBC377DB94B40 CRC64;

Query Match
Best Local Similarity 100.0%; Score 36; DB 13; Length 448;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
Db 222 LKRVHEE 228

RESULT 4
ID 073587 PRELIMINARY; PRT; 455 AA.
AC 073587;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Desmin.

```

```

GN DES.
OS Scyllorhinus stellaris.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeidae; Carcharhiniformes;
OC Scyllorhinidae; Scyllorhinus.
OX NCBI_TaxID=68454;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Schultess J., Loebbecke A., Schaffeld M., Lieb B., Markl J.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Schaffeld M., Schultess J., Loebbecke A., Lieb B., Herrmann H.,
RA Markl J.;
RT "Primary structure, expression patterns and properties of vimentin and
RT desmin in the shark Scyllorhinus stellaris.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
DR EMBL; Y15064; CAA75347.1; -
DR EMBL; AJ304374; CAC83054.1; -
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Coiled coil; Intermediate filament.
SQ SEQUENCE 455 AA; 52230 MW; B330A2FCF895BCE9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 36; DB 13; Length 455;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
Db 223 LKRVHEE 229

RESULT 5
ID 090HNS PRELIMINARY; PRT; 470 AA.
AC 090HNS;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Mutant desmin.
GN MOTDESM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99359868; PubMed=10430757;
RA Li D.X., Tapscott T., Gonzalez O., Burch P.E., Quinones M.A.,
RA Zoghbi W.A., Hill R., Bachisk L.L., Mann D.L., Roberts R.;
RT "Desmin Mutation Responsible for Dilated Cardiomyopathy.";
RL Circulation 100:461-464(1999).
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
DR EMBL; AF137053; AAF15400.1; -
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF; 1.
KW Coiled coil; Intermediate filament.
FT VARIANT
SQ SEQUENCE 470 AA; 53553 MW; 1B5D9EAE51FB319 CRC64;

Query Match
Best Local Similarity 100.0%; Score 36; DB 4; Length 470;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKRVHEE 7
Db 239 LKRVHEE 245

```

```

RESULT 6
09H319 PRELIMINARY; PRT; 470 AA.
ID 09H319;
AC 09H319;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Mutant desmin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Park K.-Y., Dalakas M.C., Semino-Mora C., Lee H.-S., Litvak S.,
RL Goldfarb L.G.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF167579; AAG41217.1;
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
FT VARIANT 406 406
SQ SEQUENCE 470 AA; 53565 MW; 1C2AEBAE4B4B319 CRC64;

Query Match 100.0%; Score 36; DB 4; Length 470;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 239 LKRVHEE 245

RESULT 7
08TD99 PRELIMINARY; PRT; 470 AA.
ID 08TD99;
AC 08TD99;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mutant desmin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dagvadorj A.;
RT "Novel desmin mutation causing complete disruption of the intermediate
RT filament network.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF487828; AAL99078.1;
SQ SEQUENCE 470 AA; 53561 MW; BB51286FE75605DF CRC64;

Query Match 100.0%; Score 36; DB 4; Length 470;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 239 LKRVHEE 245

RESULT 8
08TCX2 PRELIMINARY; PRT; 470 AA.
ID 08TCX2;
AC 08TCX2;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mutant desmin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dagvadorj A.;
RT "Proline substitution in mutant desmin breaks the alpha-helical
RT conformation and results in severe phenotype.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY083345; AAL99215.1;
SQ SEQUENCE 470 AA; 53519 MW; 06B5F052D6FE25C CRC64;

Query Match 100.0%; Score 36; DB 4; Length 470;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 239 LKRVHEE 245

RESULT 9
08TAC4 PRELIMINARY; PRT; 470 AA.
ID 08TAC4;
AC 08TAC4;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mutant desmin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21523003; PubMed=11668632;
RA Goudreau B., Dagvadorj A., Rodrigues-Lima F., Nedeliec P.,
RA Casteras-Simon M., Perret E., Langlois S., Goldfarb L., Vicart P.;
RT "Structural and functional analysis of a new desmin variant causing
RT desmin-related myopathy.";
RL Hum. Mutat. 18:386-396(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Goudreau B., Dagvadorj A., Rodrigues-Lima F., Nedeliec P.,
RA Casteras-Simon M., Perret E., Langlois S., Goldfarb L., Vicart P.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF486807; AAL93205.1;
SQ SEQUENCE 470 AA; 53504 MW; 1B5C44024DE11869 CRC64;

Query Match 100.0%; Score 36; DB 4; Length 470;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LKRVHEE 7
DB 239 LKRVHEE 245

RESULT 10
018665 PRELIMINARY; PRT; 410 AA.
ID 018665;
AC 018665;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE C47D12.5 protein.
GN C47D12.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBL_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajadasy S.;

```

```

RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "genome sequence of the nematode C.elegans: A platform for
  investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z69902; CAA93763.1; -.
SQ SEQUENCE 410 AA; 46233 MW; 3566D13571AF2423 CRC64;

Query Match
Best Local Similarity 97.2%; Score 35; DB 5; Length 410;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
Db 217 LKKIHEE 223

RESULT 11
O8UWF1 PRELIMINARY; PRT; 459 AA.
ID O8UWF1;
AC O8UWF1;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Desmin (Fragment).
GN DES.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RX NCBI_TaxID-8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-MUSCLE;
RA Scheffeld M., Schultess J., Loebbecke A., Lieb B., Herrmann H.,
  Markl J.;
RT "Primary structure, expression patterns and properties of vimentin and
  desmin in the shark Scyliorhinus stellaris.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ304373; CAC83053.1; -.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament. 1.
DR PROSITE: PS00226; IF; UNKNOWN_1.
FT NON-TER
FT NON-TER
SQ SEQUENCE 459 AA; 52098 MW; 586FBCF0C8C2E8B CRC64;

Query Match
Best Local Similarity 97.2%; Score 35; DB 13; Length 459;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
Db 224 LKKIHEE 230

RESULT 12
O48724 PRELIMINARY; PRT; 807 AA.
ID O48724;
AC O48724;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE A2928570 protein.
GN A2928570.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
RX NCBI_TaxID-3702;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, COLUMBIA;
RX MEDLINE-20083487; PubMed-10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bentle M.-I., Town C.D.,
  Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
  Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
  Cronin L.A., Shen M., Vanden S.E., Unayam L., Tallon L.J., Gill J.E.,
  Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
  Ralston J.E., Fraser C.M., Ventner J.C.;
RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
  Salzberg S.L., Fraser C.M., Ventner J.C.;
RT "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL Nature 402:761-768(1999).
DR EMBL: AC002505; AAC14505.1; -.
SQ SEQUENCE 807 AA; 89294 MW; F83F0655E1371472 CRC64;

Query Match
Best Local Similarity 97.2%; Score 35; DB 10; Length 807;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
Db 205 LKKIHEE 211

RESULT 13
O91377 PRELIMINARY; PRT; 60 AA.
ID O91377;
AC O91377;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Myosin heavy chain (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
RX NCBI_TaxID-9031;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-94227840; PubMed-8173588;
RA Canoretta Mercader B., Dizon E., Jakovic S., Zak R.;
RT "Differential expression of ventricular-like myosin heavy chain mRNA
  in developing and regenerating avian skeletal muscles.";
RL Cell. Mol. Biol. Res. 39:425-437(1993).
DR EMBL: S70443; AAB30704.1; -.
DR InterPro: IPR002928; Myosin_tail.
DR Pfam: PF01576; Myosin_tail; 1.
FT NON-TER
FT NON-TER
SQ SEQUENCE 60 AA; 7060 MW; 3165E04EAE24AF1 CRC64;

Query Match
Best Local Similarity 94.4%; Score 34; DB 13; Length 60;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
Db 53 MKKVHEE 59

RESULT 14
O91355 PRELIMINARY; PRT; 68 AA.
ID O91355;
AC O91355;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Ventricular myosin heavy chain 1 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
RX NCBI_TaxID-9031;

```

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92038448; PubMed-1936571;
 RA Bisaha J.G., Bader D.;
 RT "Identification and characterization of a ventricular-specific avian
 RT myosin heavy chain, VMHC1: expression in differentiating cardiac and
 RT skeletal muscle."
 RL Dev. Biol. 148:355-364(1991).
 DR EMBL: S64689; AAB20271.1; -
 DR InterPro: IPR002928; Myosin_tail.
 PFam: PF01576; Myosin_tail; 1.
 FT NON_TER
 SQ SEQUENCE 68 AA: 8013 MW: 8565E1CBACB4E47 CRC64;

Query Match 94.4%; Score 34; DB 13; Length 68;
 Best Local Similarity 85.7%; Pred. No. 9.9;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
 :|||||
 DB 61 MKKVHEE 67

RESULT 15

Q9IBD4 PRELIMINARY; PRT; 1937 AA.
 ID Q9IBD4
 AC Q9IBD4
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Myosin heavy chain.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus
 NX NCBT_taxid-9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WHITE LEGHORN; TISSUE-HEART;
 RX MEDLINE-20149854; PubMed-10684978;
 RA Machida S., Noda S., Furutani Y., Takao A., Momma K., Matsuoaka R.;
 RT "Complete sequence and characterization of chick ventricular myosin
 RT heavy chain in the developing atria."
 RL Blochm. Biophys. Acta 1490:333-341(2000).
 DR EMBL: AB032197; BAA92710.1; -
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR000048; IQ_region.
 DR InterPro: IPR001609; myosin_head.
 DR InterPro: IPR004009; myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR000533; Tropomyosin.
 PFam: PF00612; IQ; 2.
 DR PFam: PF00063; myosin_head; 1.
 DR PFam: PF02736; Myosin_N; 1.
 DR PFam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRINTS: PR00194; TROPOMYOSIN.
 DR Prodom: PP000355; myosin_head; 1.
 DR SMART: SM00015; IQ; 1.
 DR SMART: SM00242; MYSC; 1.
 SQ SEQUENCE 1937 AA: 223473 MW: 95A94F69CD836781 CRC64;

Query Match 94.4%; Score 34; DB 13; Length 1937;
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LKKVHEE 7
 :|||||
 DB 1930 MKKVHEE 1936

Search completed: November 13, 2002, 13:16:05
 Job time : 23.4894 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 34.7021 Seconds
(without alignments)
26.879 Million cell updates/sec

Title: US-09-856-086-3

Perfect score: 32
Sequence: 1 EALERQL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

1: /SID2/gcgdata/geneseq/emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	32	100.0	7	21	AA05927	Test antigen #3 fo
2	32	100.0	7	21	AA05929	Test antigen #5 fo
3	32	100.0	103	21	AA02130	Human secreted pro
4	32	100.0	441	19	AA20612	Human neurofilamen
5	32	100.0	543	22	AA93466	Human polypeptide,
6	30	93.8	183	23	AA682610	S. epidemidis ope
7	30	93.8	229	23	ABP40184	Streptococcus epi
8	30	93.8	384	23	ABP26811	Rice OsMA57 prote
9	29	90.6	249	20	AAW84049	Plant flowering ti
10	29	90.6	249	21	AA014277	

11	29	90.6	378	22	ABR64054	Drosophila melanog
12	29	90.6	382	14	AA333366	Sequence of purine
13	29	90.6	776	22	ABR61082	Drosophila melanog
14	29	90.6	1504	22	ABR60358	Drosophila melanog
15	29	90.6	2439	22	ABR62198	Drosophila melanog
16	29	90.6	3111	22	ABR60327	Drosophila melanog
17	28	87.5	29	21	AA080357	Amino acid sequenc
18	28	87.5	29	21	AA195951	Synthetic colled-c
19	28	87.5	29	22	AA81490	Leucine zipper o11
20	28	87.5	114	20	AA37886	Amino acid sequenc
21	28	87.5	145	22	ABR52797	Escherichia coli p
22	28	87.5	222	20	AA32125	Human LYST interac
23	28	87.5	227	22	ABR52676	Escherichia coli p
24	28	87.5	228	22	ABP28099	Streptococcus poly
25	28	87.5	258	23	AA80003	Human protein, SEQ
26	28	87.5	278	21	AA628823	Arabidopsis thalia
27	28	87.5	278	21	AA628821	Arabidopsis thalia
28	28	87.5	307	22	AA366514	Human FLEXRT-36 pr
29	28	87.5	307	23	ABR05658	Human brain derive
30	28	87.5	325	21	AA628822	Arabidopsis thalia
31	28	87.5	325	21	AA628822	Arabidopsis thalia
32	28	87.5	329	21	AA628821	Arabidopsis thalia
33	28	87.5	332	22	AA79019	Human protein, SEQ
34	28	87.5	339	20	AA15825	Hydroxyphenyl pyrro
35	28	87.5	409	23	AA668287	Modified IGP2A-1 p
36	28	87.5	457	23	AA668286	Brevibacillus late
37	28	87.5	495	22	ABR64646	Drosophila melanog
38	28	87.5	499	22	ABR66490	Drosophila melanog
39	28	87.5	638	22	ABR64380	Drosophila melanog
40	28	87.5	652	21	AA95685	Cosmid cHRIM5 enco
41	28	87.5	801	23	ABR25424	Streptococcus poly
42	28	87.5	1013	17	AA99642	Poly(ADP-ribose) p
43	28	87.5	1014	20	AA33659	Human poly(ADP-ri
44	28	87.5	1014	21	AA58043	Human poly(ADP-ri
45	28	87.5	1014	21	AA49939	Human nuclear NAD+

ALIGNMENTS

RESULT 1
ID AAB05927 standard; peptide: 7 AA.
XX AAB05927;
XX 17-OCT-2000 (first entry)
DE Test antigen #3 for spongiform and demyelinating disease diagnosis.
XX Human: cow; myelin; myelin neurofilament; immunogen; antigen;
XX bovine spongiform encephalopathy; BSE; multiple sclerosis;
KW Creutzfeldt-Jacob disease; CJD; demyelinating disease; diagnostic test.
XX Bos taurus.
OS Homo sapiens.
XX WO200031545-A1.
XX 02-JUN-2000.
XX 25-NOV-1999; 99WO-GB03936.
XX 26-NOV-1998; 98GB-0025948.
PA (UNLO) KING'S COLLEGE.
PI Edlinger A;
XX WPI; 2000-400194/34.
XX Diagnosing spongiform or demyelinating disease in vertebrates such as
PT bovine spongiform encephalopathy and Creutzfeldt-Jacob disease comprises

PT assaying a biological sample for myelin and/or myelin neurofilament
 PT antibodies -
 XX
 PS Claim 5; Page 2; 16pp; English.
 CC
 CC The present peptide may be used as a test antigen in a kit for diagnosing
 CC spongiform or demyelinating disease in vertebrates, including bovine
 CC Creutzfeldt-Jacob disease (CJD). Peptides used in the kit are antigenic
 CC components of myelin or myelin neurofilaments. Biological samples
 CC are assayed for antibodies, especially IGA antibodies, which bind to
 CC myelin and/or myelin neurofilaments or an antigenic peptide, such as the
 CC present sequence. Any reading in excess of two standard deviations of
 CC the healthy controls would indicate a positive response.
 XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 32; DB 21; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EALEKQL 7
 Db 1 EALEKQL 7
 RESULT 2
 AAB05929
 ID AAB05929 standard; peptide; 7 AA.
 XX
 AC AAB05929;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Test antigen #5 for spongiform and demyelinating disease diagnosis.
 XX
 KM Human; cow; myelin; myelin neurofilament; immunogen; antigen;
 KM bovine spongiform encephalopathy; BSE; multiple sclerosis;
 KM Creutzfeldt-Jacob disease; CJD; demyelinating disease; diagnostic test.
 XX
 OS Bos taurus.
 OS Homo sapiens.
 XX
 PM WO200031545-A1.
 XX
 PD 02-JUN-2000.
 XX
 PE 25-NOV-1999; 99WO-CB03936.
 XX
 PR 26-NOV-1998; 98GB-0035948.
 XX
 PA (UNLO) KING'S COLLEGE.
 XX
 PI Edlinger A;
 XX
 DR WPI: 2000-400194/34.
 XX
 PT Diagnosing spongiform or demyelinating disease in vertebrates such as
 PT bovine spongiform encephalopathy and Creutzfeldt-Jacob disease comprises
 PT assaying a biological sample for myelin and/or myelin neurofilament
 PT antibodies -
 XX
 PS Claim 5; Page 2; 16pp; English.
 CC
 CC The present peptide may be used as a test antigen in a kit for diagnosing
 CC spongiform or demyelinating disease in vertebrates, including bovine
 CC Creutzfeldt-Jacob disease (CJD). Peptides used in the kit are antigenic
 CC components of myelin or myelin neurofilaments. Biological samples
 CC are assayed for antibodies, especially IGA antibodies, which bind to
 CC myelin and/or myelin neurofilaments or an antigenic peptide, such as the
 CC present sequence. Any reading in excess of two standard deviations of
 CC the healthy controls would indicate a positive response.

XX
 SQ Sequence 7 AA;
 Query Match 100.0%; Score 32; DB 21; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EALEKQL 7
 Db 1 EALEKQL 7
 RESULT 3
 AAG02130
 ID AAG02130 standard; Protein; 103 AA.
 XX
 AC AAG02130;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 6211.
 XX
 KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 OS
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PE 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GSEST) GENSET.
 XX
 PI Dumas Mline Edwards J, Duclert A, Giordano J;
 XX
 DR WPI: 2000-500381/45.
 DR N-PSDB; AAC02136.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 6211; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 103 AA;
 Query Match 100.0%; Score 32; DB 21; Length 103;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EALEKQL 7
 Db 81 EALEKQL 87

RESULT 4
AAV20612 standard; Protein: 441 AA.
XX
AC AAV20612:
XX
DT 22-JUL-1999 (first entry)
XX
DE Human neurofilament-L wild type protein fragment 2.
XX
KM Human: beta-amyloid precursor protein; beta-APP; diagnosis: cancer;
KM frameshift mutation; age-related disease; neurodegenerative disorder;
KM Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KM Huntington's disease; multiple sclerosis; alcoholic liver disease;
KM diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KM ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KM neurofilament-F; presenilin I; presenilin II; cellular tumor antigen;
KM glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KM bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KM high mobility group protein-C; neuroendocrine specific protein A.
XX
OS Homo sapiens.
XX
PN WO9845322-A2.
XX
PD 15-OCT-1998.
XX
PF 02-APR-1998; 98WO-1B00705.
XX
PR 10-APR-1997; 97US-0043163.
XX
PA (UYUT-) RIJKSUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
PI Burdach JPH, Grosveld FG, Van Leeuwen FW;
XX WPI: 1998-609901/51.
DR N-PSDB; AAX5758.
XX
PT Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA
XX
PS Disclosure: Figure 7; 258bp; English.
XX
CC This invention describes a novel method for the diagnosis of a disease
CC caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumor antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
SQ Sequence 441 AA:

Query Match 100.0%; Score 32; DB 19; Length 441;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
|||
Db 386 EALEKOL 392
RESULT 5
AAM93466
ID AAM93466 standard; Protein: 543 AA.
XX
AC AAM93466:
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 3132.
XX
KM Human: full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI: 2001-524255/58.
DR N-PSDB; AAK94387.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 3132; 1380bp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5' - and 3' - ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 543 AA:

Query Match 100.0%; Score 32; DB 22; Length 543;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
|||
Db 327 EALEKOL 333

RESULT 6
AAG82610
ID AAG82610 standard; Protein: 183 AA.
XX
AC AAG82610:
XX
DT 03-SEP-2001 (first entry)

```

XX  S. epidermidis open reading frame protein sequence SEQ ID NO:2314.
DE
XX  Staphylococcus epidermidis SRI strain; infection; diagnosis;
KM
XX  vaccination; endocarditis.
OS
XX  Staphylococcus epidermidis.
PN
XX  WO200134809-A2.
PD
XX  17-MAY-2001.
XX
XX  09-NOV-2000; 2000MO-US30782.
XX
XX  09-NOV-1999; 99US-0164258.
XX
XX  (GLAX ) GLAXO GROUP LTD.
XX
XX  Kimmery WJ;
PI
XX  WPI: 2001-316495/33.
DR
XX  N-PSDB; AAH53460.
XX
XX  Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT  useful for vaccinating against infections, e.g. endocarditis -
XX
XX  Claim 18: Page 615; 2188pp; English.
PS
XX  AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC  (II), given in AG81454 to AG83120, from Staphylococcus epidermidis.
CC  (I) and (II) can have antibacterial activity and therefore can be used
CC  in vaccination. The nucleic acids (I) may be used to produce the
CC  S. epidermidis polypeptides (II) via the production of vectors
CC  containing them which are used to produce hosts cells which express the
CC  polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC  used to vaccinate subjects and to raise antibodies against the bacteria.
CC  The polypeptides may also be used to assay for other inhibitors of their
CC  activity and therefore identify compounds that may be used for the
CC  treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC  AAH55000 represent specifically claimed S. epidermidis genomic DNA
CC  polynucleotide sequences from the present invention. AAH5091 to
CC  AAH55098 represent oligonucleotide sequences and primers which are used
CC  in the exemplification of the present invention.
CC  N.B. The present invention specifically claims all the polynucleotide
CC  sequences given in the sequence listing of the present specification,
CC  however the sequence listing only goes up to SEQ ID NO:4454 so even
CC  though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC  no sequences are present for SEQ ID NO:4455 to 4464.
XX
XX  Sequence 183 AA;
SQ
XX
XX  Query Match 93.8%; Score 30; DB 22; Length 183;
XX  Best Local Similarity 85.7%; Pred. No. 3e+02;
XX  Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EALEKQL 7
Db 97 EALEKQL 103

```

```

OS  Staphylococcus epidermidis.
XX
XX  US6380370-B1.
PN
XX  30-APR-2002.
PD
XX
XX  13-AUG-1998; 98US-0134001.
XX
XX  14-AUG-1997; 97US-055779P.
PR  08-NOV-1997; 97US-064964P.
XX
XX  (GENO-) GENOME THERAPEUTICS CORP.
PA
XX  Doucette-Stamm LA, Bush D;
PI
XX  WPI: 2002-381255/41.
DR
XX  N-PSDB; ABN92729.
XX
XX  Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT  polypeptide, useful for diagnosing and treating bacterial infections -
XX
XX  Disclosure; SEQ ID 5029; 267pp; English.
XX
XX  ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
CC  frame (ORF) nucleic acid sequences which encode the amino acid sequences
CC  given in ABP35124 to ABP37960. The S. epidermidis sequences have
CC  antibacterial activity and can be used in gene therapy. The sequences
CC  can also be used in the diagnosis and treatment of bacterial infections,
CC  particularly S. epidermidis infections. The sequences can be used to
CC  screen for compounds able to interfere with the S. epidermidis life
CC  cycle or inhibit S. epidermidis infection.
CC  N.B. The sequence data for this patent did not form part of the printed
CC  specification, but was obtained in electronic format directly from the
CC  USPO web site.
XX
XX  Sequence 229 AA;
SQ
XX
XX  Query Match 93.8%; Score 30; DB 23; Length 229;
XX  Best Local Similarity 85.7%; Pred. No. 3.7e+02;
XX  Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 EALEKQL 7
Db 105 EALEKQL 111

```

```

RESULT 8
ABP26811
ID  ABP26811 standard; Protein; 384 AA.
XX
XX  ABP26811;
AC
XX
XX  02-JUL-2002 (first entry)
DT
XX
XX  Streptococcus polypeptide SEQ ID NO 2798.
DE
XX  Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM  group A streptococcus; Streptococcus pyogenes; antibacterial;
XX  antiinflammatory; infection; vaccine; meningitis; gene therapy.
OS  Streptococcus agalactiae.
PN
XX  WO200234771-A2.
PD
XX
XX  02-MAY-2002.
XX
XX  29-OCT-2001; 2001WO-GB04789.
XX
XX  27-OCT-2000; 2000GB-0026333.
PR  24-NOV-2000; 2000GB-0028727.
XX  07-MAR-2001; 2001GB-0005640.
XX
XX  (CHIR-) CHIRON SPA.
PA

```

PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 DR WPI: 2002-352536/38.
 DR N-PSDB: ABN67442.
 XX
 PT New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 3431; 4525pp: English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (i), nucleic acids encoding (i), ABN66044-ABN71526 and
 CC antibodies that bind (i) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (i) are used to detect Streptococcus in a
 CC biological sample. (i) is used to determine whether a compound binds to
 CC (i). A composition comprising (i) or a nucleic acid encoding (i), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (i) may be used to recombinantly produce (i) and may be
 CC used in gene therapy. Antibodies to (i) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 CC
 SQ Sequence 384 AA;
 XX
 QY 1 EALEKOL 7
 DB 190 EALEKOL 196
 XX
 Query Match 93.8%; Score 30; DB 23; Length 384;
 Best Local Similarity 85.7%; Pred. No. 6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 9
 ID AAW84049 standard; Protein; 249 AA.
 XX
 AC AAW84049;
 XX
 DT 15-MAR-1999 (first entry)
 XX
 DE Rice OSMAD57 protein involved in floral development.
 XX
 KM OSMAD57; rice; flower development; apical dominance; dwarfism;
 KM transgenic plant.
 XX
 OS Oryza sativa.
 XX
 FH Key Location/Qualifiers
 FT Domain 2..57
 FT /Label= MADS-box
 FT Domain 95..159
 FT /Label= K-box
 XX
 PN WO9854328-A2.
 XX
 PD 03-DEC-1998.
 XX
 PF 02-JUN-1998; 98WO-US11278.
 XX
 PR 02-JUN-1997; 97OS-0867087.
 XX
 PA (UNIV) UNIV WASHINGTON STATE RES FOUND.

XX
 PI An G;
 XX
 DR WPI: 1999-035176/03.
 DR N-PSDB: AAV71740.
 XX
 PT New plant MADS-box genes - useful for, e.g. controlling floral
 PT development and apical dominance
 XX
 PS Example 1; Fig 2; 53pp: English.
 XX
 CC This is the amino acid sequence of rice OSMAD57, as deduced from an
 CC isolated cDNA clone (see AAV71740). OSMAD7 is a novel MADS-box gene
 CC involved in the control of floral development and apical dominance.
 CC OSMAD56 (see AAW84048) and OSMAD58 (see AAW84050) proteins are also
 CC disclosed. Expression of OSMAD56, OSMAD57 or OSMAD58 in a
 CC transgenic plant causes the plant to exhibit at least one phenotype
 CC selected from: (i) diminished apical dominance, (ii) early
 CC flowering, (iii) altered daylength requirement for flowering, (iv)
 CC greater synchronisation of flowering, and (v) relaxed vernalisation
 CC requirement, compared to a non-transgenic control plant. Claimed
 CC methods of producing such a transgenic plant utilise: (i) OSMAD56,
 CC OSMAD57 or OSMAD58 nucleic acids (see AAV71739-41).
 CC
 SQ Sequence 249 AA;
 XX
 QY 1 EALEKOL 7
 DB 128 ESEKOL 134
 XX
 Query Match 90.6%; Score 29; DB 20; Length 249;
 Best Local Similarity 85.7%; Pred. No. 6.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 10
 ID AAO14277 standard; Protein; 249 AA.
 XX
 AC AAO14277;
 XX
 DT 05-AUG-2002 (first entry)
 XX
 DE Plant flowering time control related protein #3.
 XX
 KM Plant; flowering time.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 228..229
 FT /note= "encoded by CCM"
 FT MISC-difference 249
 FT /note= "encoded by CTRACA"
 XX
 PN KR99030639-A.
 XX
 PD 06-MAY-1999.
 XX
 PF 02-OCT-1997; 97KR-0050938.
 XX
 PR 02-OCT-1997; 97KR-0050938.
 XX
 PA (UNIV) UNIV POHANG SCI & TECHNOLOGY.
 XX
 PI Ahn GS, Kang HG, Seong SG, Ahn JH;
 XX
 DR WPI: 2000-335956/29.
 DR N-PSDB: AAL44854.
 XX
 PT Gene controlling the flowering time, vector containing the same and
 PT plants transformed thereby -
 XX

PS Disclosure; Fig 7; 36pp; Korean.

CC The present invention relates to genes capable of controlling the
CC flowering time of plants. The present sequence is a protein
CC described in the exemplification of the invention.

XX
SQ Sequence 249 AA;

Query Match 90.6%; Score 29; DB 21; Length 249;
Best Local Similarity 85.7%; Pred. No. 6.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
Db 128 ESIKOL 134

RESULT 11

ID ABB64054 standard; Protein; 378 AA.

XX ABB64054;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 18954.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EW;

XX MPI; 2001-656860/75.

XX N-PSDB; ABL08157.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.

XX Disclosure: SEQ ID NO 18954; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
XX sequences (ABLI6176-ABLI30511) and the encoded proteins

XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 378 AA;

Query Match 90.6%; Score 29; DB 22; Length 378;
Best Local Similarity 85.7%; Pred. No. 9.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7

Db 271 EAVEKOL 277

RESULT 12

ID AAR33366 standard; Protein; 382 AA.

XX AAR33366;

XX 30-JUN-1993 (first entry)

XX Sequence of purine-rich repeat (GA repeat) binding protein

XX (GABP) subunit beta 1.

XX GA binding protein; cis-regulatory element;

XX VP16 mediated induction.

XX Mus musculus.

XX Key Location/Qualifiers

XX Misc-difference 333..382

XX /note="not found in GABP-beta-2"

XX WO9304166-A.

XX 04-MAR-1993.

XX 17-AUG-1992; 92WO-US06748.

XX 16-AUG-1991; 91US-0746032.

XX (CARN-) CARNEGIE INST WASHINGTON.

XX Lamarco KL, Mc Knight SL, Thompson CC;

XX MPI; 1993-093998/11.

XX N-PSDB; AAQ37966.

XX DNA encoding GA binding protein sub-unit - allows investigation

XX of sub-unit sequence motif functions, for control of rapid cell

XX division e.g. in cancer

XX Disclosure: Fig 2B1-3; 68pp; English.

XX A cis-regulatory element required for virion associated protein VP16
XX mediated induction of herpes simplex virus 1 (HSV1) immediate early
XX (IE) genes consists of three imperfect repeats of the purine-rich
XX hexanucleotide 5'-CGGAR-3'. A protein complex capable of avid
XX interaction with the purine-rich repeats (GA repeats) has been
XX identified in soluble preparations of rat liver nucleic. This GA
XX binding protein (GABP) consists of two separable subunits.
XX Applicants have isolated cDNA clones encoding both subunits of GABP
XX and have revealed that one (GABP alpha) is related to the Ets
XX transforming protein, while the other (GABP beta) contains a
XX series of 33-amino acid repeats related in sequence to a variety
XX of proteins. The sequences for GABP-beta-1 and GABP-beta-2 are
XX identical up to nucleotide 1130 except for a three nucleotide
XX insertion (GTA) at posn. 828 of GABP-beta-1.

XX Sequence 382 AA;

Query Match 90.6%; Score 29; DB 14; Length 382;
Best Local Similarity 85.7%; Pred. No. 9.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
Db 335 EALOKOL 341

RESULT 13

ABB61082

ID ABB61082 standard; Protein; 776 AA.
XX
AC ABB61082;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 10038.
XX
KM Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PT 11-JUL-2000; 2000US-0614150.
XX
PS (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL05185.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Disclosure; SEQ ID NO 10038; 21pp + Sequence listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
XX sequences (ABU1840-ABU16175) and the encoded proteins
XX (ABU37737-ABU72072).
XX
CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 776 AA;
XX
Query Match 90.6%; Score 29; DB 22; Length 776;
XX Best Local Similarity 85.7%; Pred. No. 1.8e+03;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 EALEKOL 7
XX
DB 203 EAVEKOL 209
XX
RESULT 14
XX ABB60358
ID ABB60358 standard; Protein; 1504 AA.
XX
AC ABB60358;
XX
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 7866.
XX
KM Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;

PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PT 11-JUL-2000; 2000US-0614150.
XX
PS (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL04461.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
PS Disclosure; SEQ ID NO 7866; 21pp + Sequence listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
XX sequences (ABU1840-ABU16175) and the encoded proteins
XX (ABU37737-ABU72072).
XX
CC The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1504 AA;
XX
Query Match 90.6%; Score 29; DB 22; Length 1504;
XX Best Local Similarity 85.7%; Pred. No. 3.3e+03;
XX Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 EALEKOL 7
XX
DB 633 EAVEKOL 639
XX
RESULT 15
XX ABB62198
ID ABB62198 standard; Protein; 2439 AA.
XX
AC ABB62198;
XX
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 13386.
XX
KM Drosophila: developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PT 11-JUL-2000; 2000US-0614150.
XX
PS (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI: 2001-656860/75.
DR N-PSDB: ABL06301.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions .
XX
PS Disclosure: SEQ ID NO 13386; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABH57737-ABH72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2439 AA;

Query Match 90.6%; Score 29; DB 22; Length 2439;
Best Local Similarity 85.7%; Pred. No. 5.2e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKQL 7
|||:|
Db 1964 EALEKEL 1970

Search completed: November 13, 2002, 13:22:07
Job time : 35.7021 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 10.1277 Seconds
(without alignments)
20.336 Million cell updates/sec

Title: US-09-856-086-3
Perfect score: 32
Sequence: 1 EALEKQL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.CONB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.CONB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.CONB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.CONB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTOTUS.CONB.pep:*
6: /cgn2_6/ptodata/1/1aa/Dackfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	160	2	US-08-726-306A-183 Sequence 183, App
2	30	93.8	229	4	US-09-134-001C-5029 Sequence 5029, Ap
3	29	90.6	249	2	US-08-867-087B-15 Sequence 15, Appl
4	28	87.5	1013	4	US-08-860-886-2 Sequence 2, Appl
5	28	87.5	1014	4	US-09-078-347A-3 Sequence 3, Appl
6	27	84.4	14	2	US-08-491-527A-8 Sequence 8, Appl
7	27	84.4	94	6	5218099-7 Patent No. 5218099
8	27	84.4	248	6	US-08-867-087B-17 Sequence 17, Appl
9	27	84.4	853	2	US-08-468-558-3 Sequence 3, Appl
10	26	84.4	853	4	US-08-676-444-3 Sequence 3, Appl
11	26	81.2	45	4	US-08-960-190A-33 Sequence 33, Appl
12	26	81.2	45	4	US-09-695-458-26 Sequence 26, Appl
13	26	81.2	64	4	US-09-695-458-31 Sequence 31, Appl
14	26	81.2	65	2	US-08-867-087B-32 Sequence 32, Appl
15	26	81.2	65	2	US-08-867-087B-33 Sequence 33, Appl
16	26	81.2	65	2	US-08-867-087B-34 Sequence 34, Appl
17	26	81.2	128	2	US-08-370-909-2 Sequence 2, Appl
18	26	81.2	241	2	US-08-867-087B-11 Sequence 11, Appl
19	26	81.2	285	6	5171685-4 Patent No. 5171685
20	26	81.2	285	6	5171685-4 Patent No. 5171685
21	26	81.2	293	4	US-09-695-458-23 Sequence 23, Appl
22	26	81.2	318	4	US-09-695-458-22 Sequence 22, Appl
23	26	81.2	353	4	US-09-184-970B-171 Sequence 171, App
24	26	81.2	434	3	US-08-444-644-23 Sequence 23, Appl
25	26	81.2	434	3	US-08-444-644-31 Sequence 31, Appl
26	26	81.2	434	3	US-08-444-644-40 Sequence 40, Appl
27	26	81.2	434	3	US-08-444-644-46 Sequence 46, Appl

28	26	81.2	434	4	US-08-232-246A-23 Sequence 23, Appl
29	26	81.2	434	4	US-08-232-246A-31 Sequence 31, Appl
30	26	81.2	434	4	US-08-232-246A-40 Sequence 40, Appl
31	26	81.2	434	4	US-08-232-246A-46 Sequence 46, Appl
32	26	81.2	509	2	US-08-737-524B-27 Sequence 27, Appl
33	26	81.2	535	4	US-09-134-001C-3338 Sequence 3338, Ap
34	26	81.2	1220	3	US-08-930-996A-2 Sequence 2, Appl
35	26	81.2	1285	1	US-07-582-945-2 Sequence 2, Appl
36	26	81.2	1285	2	US-08-453-141-2 Sequence 2, Appl
37	26	81.2	1285	3	US-08-293-314-2 Sequence 2, Appl
38	25	78.1	22	3	US-08-940-095-24 Sequence 24, Appl
39	25	78.1	22	3	US-08-940-095-40 Sequence 40, Appl
40	25	78.1	22	3	US-08-940-095-66 Sequence 66, Appl
41	25	78.1	22	3	US-08-940-095-80 Sequence 80, Appl
42	25	78.1	22	3	US-08-940-095-129 Sequence 129, App
43	25	78.1	22	3	US-08-940-095-132 Sequence 132, App
44	25	78.1	22	3	US-08-940-095-161 Sequence 161, App
45	25	78.1	22	3	US-08-940-095-178 Sequence 178, App

ALIGNMENTS

```
RESULT 1
US-08-726-306A-183
; Sequence 183, Application US/08726306A
; Patent No. 5958684
;
GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Burbach, Johannes Peter Henri
; APPLICANT: Grosveld, Franklin G.
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
; FILING DATE: 02-Oct-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; TELECOMMUNICATION INFORMATION:
; INFORMATION FOR SEQ ID NO: 183:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 160 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
;
US-08-726-306A-183
;
Query Match 100.0%; Score 32; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 1 EALEKOL 7
1111111
Db 113 EALEKOL 119

RESULT 2

US-09-134-001C-5029
Sequence 5029, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5029
LENGTH: 229
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5029

Query Match 93.8%; Score 30; DB 4; Length 229;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
1111111
Db 105 EALEKOL 111

RESULT 3

US-08-867-087B-15
Sequence 15, Application US/08867087B
Patent No. 5990386
GENERAL INFORMATION:
APPLICANT: An, Gynheung
TITLE OF INVENTION: GENES CONTROLLING FLORAL DEVELOPMENT
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/08/867,087B
CURRENT FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5029
LENGTH: 229
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-08-867-087B-15

Query Match 93.8%; Score 30; DB 4; Length 229;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

REFERENCE/DOCKET NUMBER: 4630-47071

TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acid residues
TYPE: amino acid
TOPOLOGY: linear
US-08-867-087B-15

Query Match 90.6%; Score 29; DB 2; Length 249;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
1111111
Db 128 EALEKOL 134

RESULT 4

US-08-860-886-2
Sequence 2, Application US/08860886
Patent No. 6335009
GENERAL INFORMATION:
APPLICANT: Burkle, Alexander
APPLICANT: Zur Hausen, Harald
APPLICANT: Jan-Helner, Kupper
TITLE OF INVENTION: VECTORS AND VIRUSES FOR USE
IN GENE THERAPY
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/08/860,886
CURRENT FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5029
LENGTH: 229
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-08-860-886-2

Query Match 87.5%; Score 28; DB 4; Length 1013;
Best Local Similarity 85.7%; Pred. No. 7.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 193 EALKKOL 199

|||||

RESULT 5

US-09-078-347A-3

Sequence 3, Application US/09078347A

Patent No. 6133968

GENERAL INFORMATION:

APPLICANT: Le, Xiao-Chun

APPLICANT: Weinfield, Michael

TITLE OF INVENTION: Methods for Quantitating Low Level

TITLE OF INVENTION: Modifications of Nucleotide Sequences

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Medien & Carroll, LLP

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/078,347A

FILING DATE: 13-MAY-1998

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Carroll, Peter G.

REGISTRATION NUMBER: 32,837

REFERENCE/DOCKET NUMBER: DALB-03283

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1014 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-09-078-347A-3

Query Match 87.5%; Score 28; DB 4; Length 1014;

Best Local Similarity 85.7%; Pred. No. 7, 5e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKKOL 7

Db 193 EALKKOL 199

RESULT 6

US-08-491-527A-8

Sequence 8, Application US/08491527A

Patent No. 5824483

GENERAL INFORMATION:

APPLICANT: Houston, Michael E.

APPLICANT: Hodges, Robert S.

TITLE OF INVENTION: Conformationally-Restricted Combinatorial

TITLE OF INVENTION: Library Composition and Method

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: 350 Cambridge Avenue, Suite 250

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/491,527A

FILING DATE: 16-JUNE-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/260,199

FILING DATE: 15-JUN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/245,507

FILING DATE: 18-MAY-1994

ATTORNEY/AGENT INFORMATION:

NAME: Sholtz, Charles K.

REGISTRATION NUMBER: 38,615

REFERENCE/DOCKET NUMBER: 7900-0008.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: peptide EK, linear 7

US-08-491-527A-8

Query Match 84.4%; Score 27; DB 2; Length 14;

Best Local Similarity 71.4%; Pred. No. 18;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKKOL 7

Db 3 EALEKKOL 9

RESULT 7

US-08-491-527A-8

Sequence 9, Application US/08491527A

Patent No. 5218099

GENERAL INFORMATION:

APPLICANT: REYES, GREGORY R.; BRADLEY, DANIEL W.; RABIN, LINDA;

APPLICANT: KIRK

TITLE OF INVENTION: POST-TRANSFUSION, NON-A, NON-B HEPATITIS

NUMBER OF SEQUENCES: 17

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/372,711

FILING DATE: 28-JUN-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 350,570; 334,701; 228,334; 215,728; 846,757

FILING DATE: 11-MAY-1989

APPLICATION NUMBER: 334,701

FILING DATE: 06-APR-1989

APPLICATION NUMBER: 228,334

FILING DATE: 04-AUG-1988

APPLICATION NUMBER: 215,728

FILING DATE: 06-JUL-1988

APPLICATION NUMBER: 846,757

FILING DATE: 01-APR-1986

SEQ ID NO: 7;

LENGTH: 94

Query Match 84.4%; Score 27; DB 6; Length 94;

Best Local Similarity 71.4%; Pred. No. 1.2e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
11:1111
Db 7 EAMDKOL 13

RESULT 8
US-08-867-087B-17

Sequence 17, Application US/08867087B
Patent No. 5990386

GENERAL INFORMATION:

APPLICANT: An, Gynheung
TITLE OF INVENTION: GENES CONTROLLING FLORAL DEVELOPMENT
TITLE OF INVENTION: AND APICAL DOMINANCE IN PLANTS
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:

ADDRESSEE: Klarquist Sparkman Campbell Leigh &
Whinston, LLP
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
STREET: Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: United States of America
ZIP: 97204

COMPUTER READABLE FORM:

MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/867.087B
FILING DATE: June 2, 1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/323,449
FILING DATE: October 14, 1994
APPLICATION NUMBER: U.S. 08/485,981
FILING DATE: June 7, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Dow, Alan. E.

REGISTRATION NUMBER: 35,123

REFERENCE/DOCKET NUMBER: 4630-47071

TELECOMMUNICATION INFORMATION:

TELEPHONE: (503) 226-7391

TELEFAX: (503) 228-9446

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 248 amino acid residues

TYPE: amino acid

TOPOLOGY: linear

US-08-867-087B-17

Query Match 84.4%; Score 27; DB 2; Length 248;

Best Local Similarity 85.7%; Pred. No. 3e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EALEKOL 7
111111

Db 128 EOLEKOL 134

RESULT 9
US-08-468-558-3

Sequence 3, Application US/08468558

Patent No. 5877280

GENERAL INFORMATION:

APPLICANT: Wetmur, James G.

TITLE OF INVENTION: Cloning and Expression of Thermostable

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive

CITY: Lexington
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/468,558

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: MSN94-05

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 853 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-468-558-3

Query Match 84.4%; Score 27; DB 2; Length 853;

Best Local Similarity 100.0%; Pred. No. 9.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ALEKOL 7
111111

Db 522 ALEKOL 527

RESULT 10
US-08-676-444-3

Sequence 3, Application US/08676444A

Patent No. 6294325

GENERAL INFORMATION:

APPLICANT: Wetmur, James G.

TITLE OF INVENTION: CLONING AND EXPRESSION OF THERMOSTABLE

FILE REFERENCE: MSM95-02

CURRENT APPLICATION NUMBER: US/08/676,444A

NUMBER OF SEQ ID NOS: 48

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 853

TYPE: PRT

ORGANISM: Escherichia coli

US-08-676-444-3

Query Match 84.4%; Score 27; DB 4; Length 853;

Best Local Similarity 100.0%; Pred. No. 9.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 ALEKOL 7
111111

Db 522 ALEKOL 527

RESULT 11
US-08-960-190A-33

Sequence 33, Application US/08960190A

Patent No. 6232445

GENERAL INFORMATION:

APPLICANT: Rhode, Peter R.
ADDRESSEE: Acevedo, Jorge

APPLICANT: Burkhardt, Martin
APPLICANT: Jiao, Jin-an
APPLICANT: Wong, Hing C.
TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND
NUMBER OF SEQUENCES: 38
TITLE OF INVENTION: METHODS OF USE THEREOF
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dike, Bronsteijn, Roberts & Cushman, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: usa
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,190A
FILING DATE: 29-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Corless, Peter F.
REGISTRATION NUMBER: 33,860
REFERENCE/DOCKET NUMBER: 48002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-960-190A-33

Query Match 81.2%; Score 26; DB 4; Length 45;
Best Local Similarity 71.4%; Pred. No. 88;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
Db 37 QALEKEL 43

RESULT 12
US-09-695-458-26
Sequence 26, Application US/09695458
Patent No. 6380361
GENERAL INFORMATION:
APPLICANT: Conklin, Darrell C.
APPLICANT: Taft, David W.
TITLE OF INVENTION: Educational Kit and Method Containing No. 6380361el Alpha Helical
FILE REFERENCE: 99-78
CURRENT APPLICATION NUMBER: US/09/695,458
CURRENT FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/162,623
PRIOR FILING DATE: 1999-10-29
NUMBER OF SEQ ID NOS: 41
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 26
LENGTH: 45
TYPE: PRT
ORGANISM: Homo sapiens
US-09-695-458-26

Query Match 81.2%; Score 26; DB 4; Length 45;
Best Local Similarity 71.4%; Pred. No. 88;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
Db 19 EALOKEL 25

RESULT 13
US-09-695-458-31
Sequence 31, Application US/09695458
Patent No. 6380361
GENERAL INFORMATION:
APPLICANT: Conklin, Darrell C.
APPLICANT: Taft, David W.
TITLE OF INVENTION: Educational Kit and Method Containing No. 6380361el Alpha Helical
FILE REFERENCE: 99-78
CURRENT APPLICATION NUMBER: US/09/695,458
CURRENT FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: 60/162,623
PRIOR FILING DATE: 1999-10-29
NUMBER OF SEQ ID NOS: 41
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 31
LENGTH: 64
TYPE: PRT
ORGANISM: Homo sapiens
US-09-695-458-31

Query Match 81.2%; Score 26; DB 4; Length 64;
Best Local Similarity 71.4%; Pred. No. 12e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
Db 10 EALOKEL 16

RESULT 14
US-08-867-087B-32
Sequence 32, Application US/08867087B
Patent No. 5990386
GENERAL INFORMATION:
APPLICANT: An, Gynheung
TITLE OF INVENTION: GENES CONTROLLING FLORAL DEVELOPMENT
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klargust, Sparkman Campbell Leigh &
ADDRESSEE: Whinston, LLP
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
STREET: Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: United States of America
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,087B
FILING DATE: June 2, 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/323,449
FILING DATE: October 14, 1994
APPLICATION NUMBER: U.S. 08/485,981
FILING DATE: June 7, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Dow, Alan. E.
 REGISTRATION NUMBER: 35,123
 REFERENCE/DOCKET NUMBER: 4630-47071
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (503) 226-7391
 TELEFAX: (503) 228-9446
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 65 amino acid residues
 TYPE: amino acid
 TOPOLOGY: linear
 US-08-867-087B-32

Query Match

Best Local Similarity 81.2%; Score 26; DB 2; Length 65;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7

DB 34 QALEROL 40

RESULT 15

US-08-867-087B-33
 Sequence 33, Application US/08867087B
 Patent No. 5980386

GENERAL INFORMATION:

APPLICANT: An, Gynheung
 TITLE OF INVENTION: GENES CONTROLLING FLORAL DEVELOPMENT
 TITLE OF INVENTION: AND APICAL DOMINANCE IN PLANTS
 NUMBER OF SEQUENCES: 70
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Klarquist Sparkman Campbell Leigh &
 ADDRESSEE: Whinston, LLP

STREET: One World Trade Center
 STREET: 121 S.W. Salmon Street

STREET: Suite 1600

CITY: Portland

STATE: Oregon

COUNTRY: United States of America

ZIP: 97204

COMPUTER READABLE FORM:

MEDIUM TYPE: Disk, 3-1/2 inch

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS DOS

SOFTWARE: WordPerfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/867,087B

FILING DATE: June 2, 1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. 08/323,449

FILING DATE: October 14, 1994

APPLICATION NUMBER: U.S. 08/485,981

FILING DATE: June 7, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Dow, Alan. E.

REGISTRATION NUMBER: 35,123

REFERENCE/DOCKET NUMBER: 4630-47071

TELECOMMUNICATION INFORMATION:

TELEPHONE: (503) 226-7391

TELEFAX: (503) 228-9446

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 65 amino acid residues

TYPE: amino acid

TOPOLOGY: linear

US-08-867-087B-33

QY 1 EALEKOL 7
 DB 34 EALEROL 40

Search completed: November 13, 2002, 13:18:06
 Job time: 11.1277 secs

Query Match 81.2%; Score 26; DB 2; Length 65;
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

APPLICANT: Schwartz, John
APPLICANT: Hamad, Kimberly
APPLICANT: Zhang, Shuangyong
TITLE OF INVENTION: Direct, Externally Imposed Control of Polypeptides
FILE REFERENCE: MJB-086
CURRENT APPLICATION NUMBER: US/09/905,831
PRIOR FILING DATE: 2001-07-13
PRIOR APPLICATION NUMBER: US 60/276,313
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: US 60/218,312
PRIOR FILING DATE: 2000-07-14
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn version 3.0
SEQ ID NO 15
LENGTH: 29
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: tColl-L129 trimeric bundle
US-09-905-831-15

Query Match 87.5%; Score 28; DB 10; Length 29;
Best Local Similarity 85.7%; Pred. No. 5.7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
Db 3 EALEKOL 9

RESULT 3
US-09-864-761-37915
Sequence 37915, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 37915
LENGTH: 34
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL035416.7
OTHER INFORMATION: EXPRESSED IN HB1100, SIGNAL = 0.96
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.99
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
OTHER INFORMATION: EST HUMAN HT: BF31330.1, VALUE 4.00e-07
OTHER INFORMATION: SWISSPROT HIT: P23886, VALUE 7.20e-01
US-09-864-761-37915

Query Match 84.4%; Score 27; DB 10; Length 34;
Best Local Similarity 85.7%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EALEKOL 7
Db 13 EALEKOL 19

RESULT 4
US-09-801-574-63
Sequence 63, Application US/09801574
Patent No. US20020081592A1
GENERAL INFORMATION:
APPLICANT: Wang, Peijiang Jeremy
APPLICANT: Page, David C.
TITLE OF INVENTION: Reproduction-Specific Genes
FILE REFERENCE: 0399, 2007-002
CURRENT APPLICATION NUMBER: US/09/801,574
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: 60/187,518
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: 60/261,557
PRIOR FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 90
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 63
LENGTH: 122
TYPE: PRT
ORGANISM: Homo sapiens
US-09-801-574-63

Query Match 84.4%; Score 27; DB 10; Length 122;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EALEKOL 7
Db 50 EALEKOL 56

RESULT 5
US-09-801-574-20
Sequence 20, Application US/09801574
Patent No. US20020081592A1
GENERAL INFORMATION:
APPLICANT: Wang, Peijiang Jeremy
APPLICANT: Page, David C.


```
;; TITLE OF INVENTION: Reproduction-Specific Genes
;; FILE REFERENCE: 0399.2007-002
;; CURRENT APPLICATION NUMBER: US/09/801,574
;; PRIOR APPLICATION NUMBER: 60/187,518
;; PRIOR FILING DATE: 2000-03-07
;; PRIOR APPLICATION NUMBER: 60/261,557
;; PRIOR FILING DATE: 2001-01-12
;; NUMBER OF SEQ ID NOS: 90
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 20
;; LENGTH: 123
;; TYPE: PRT
;; ORGANISM: Mus musculus
US-09-801-574-20

Query Match      84.4%; Score 27; DB 10; Length 123;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 EALEKQL 7
Db      51 EALEKDL 57

RESULT 6
US-10-052-586-540
;; Sequence 540, Application US/10052586
;; Patent No. US20020127584A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Chen, Jian
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Pan, James
;; APPLICANT: Smith, Victoria
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3430R1C1
;; CURRENT APPLICATION NUMBER: US/10/052,586
;; PRIOR FILING DATE: 2002-01-15
;; PRIOR APPLICATION NUMBER: 60/059263
;; PRIOR FILING DATE: 1997-09-18
;; PRIOR APPLICATION NUMBER: 60/059266
;; PRIOR FILING DATE: 1997-09-18
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063120
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063121
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063486
;; PRIOR FILING DATE: 1997-10-21
;; PRIOR APPLICATION NUMBER: 60/063540
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063541
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063544
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063564
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063734
;; PRIOR FILING DATE: 1997-10-29
;; PRIOR APPLICATION NUMBER: 60/063870
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/065311
```

```
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066120
;; PRIOR FILING DATE: 1997-11-21
;; PRIOR APPLICATION NUMBER: 60/066466
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/066772
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/069335
;; PRIOR FILING DATE: 1997-12-11
;; PRIOR APPLICATION NUMBER: 60/069425
;; PRIOR FILING DATE: 1997-12-12
;; PRIOR APPLICATION NUMBER: 60/069870
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/068017
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/077450
;; PRIOR FILING DATE: 1998-03-10
;; PRIOR APPLICATION NUMBER: 60/077632
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/077649
;; PRIOR FILING DATE: 1998-03-11
;; PRIOR APPLICATION NUMBER: 60/078886
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/078939
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079664
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/079786
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
```

PRIOR APPLICATION NUMBER: 60/085573
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085580
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085582
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085700
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/086023
 PRIOR FILING DATE: 1998-05-18
 PRIOR APPLICATION NUMBER: 60/086392
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/086486
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/087098
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087208
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/087609
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087759
 PRIOR FILING DATE: 1998-06-02
 PRIOR APPLICATION NUMBER: 60/087827
 PRIOR FILING DATE: 1998-06-03
 PRIOR APPLICATION NUMBER: 60/088025
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088028
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088029
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088033
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088167
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088202
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088212
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088217
 PRIOR FILING DATE: 1998-06-05
 PRIOR APPLICATION NUMBER: 60/088326
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088655
 PRIOR FILING DATE: 1998-06-09
 PRIOR APPLICATION NUMBER: 60/088722
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088738
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088740
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088811
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088824
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088825
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088826
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088863
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089090
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514

PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089908

Query Match
 Best Local Similarity 84.4%; Score 27; DB 12; Length 159;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EALEKQL 7
 Db 25 EALEKQL 31

RESULT 7
 US-09-764-864-1129
 Sequence 1129, Application US/09764864
 Patent No. US20020132753A1
 GENERAL INFORMATION:
 APPLICANT: Rosen et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 FILE REFERENCE: PT23
 CURRENT APPLICATION NUMBER: US/09/764,864
 CURRENT FILING DATE: 2001-01-17
 Prior application data removed - consult PAM or file wrapper
 NUMBER OF SEQ ID NOS: 1792
 SOFTWARE: Patentln Ver. 2.0
 SEQ ID NO 1129
 LENGTH: 160
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-764-864-1129

Query Match
 Best Local Similarity 71.4%; Score 27; DB 10; Length 160;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKQL 7
 Db 132 EALEKQL 138

RESULT 8
 US-09-801-368-32
 Sequence 32, Application US/09801368
 Patent No. US20020128250A1
 GENERAL INFORMATION:
 APPLICANT: Busby, Robert
 APPLICANT: Cali, Brian
 APPLICANT: Hecht, Peter
 APPLICANT: Holtzman, Doug
 APPLICANT: Madden, Kevin
 APPLICANT: Maxon, Mary
 APPLICANT: Milne, Todd
 APPLICANT: No. US20020128250A1man, Thea
 APPLICANT: Royer, John
 APPLICANT: Salama, Sofie
 APPLICANT: Sherman, Amit
 APPLICANT: Silva, Jeff
 APPLICANT: Summers, Eric
 TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fun
 FILE REFERENCE: 109272.147
 CURRENT APPLICATION NUMBER: US/09/801,368
 CURRENT FILING DATE: 2001-03-07
 PRIOR APPLICATION NUMBER: US 09/487,558
 PRIOR FILING DATE: 2000-01-19
 PRIOR APPLICATION NUMBER: US 60/160,587
 PRIOR FILING DATE: 1999-10-20
 NUMBER OF SEQ ID NOS: 440

;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 32
;; LENGTH: 770
;; TYPE: PRT
;; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-32

Query Match 84.4%; Score 27; DB 10; Length 770;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ALEKOL 7
DB 210 EKLEKOL 216

RESULT 9
US-09-815-242-10272

;; Sequence 10272, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 10272
;; LENGTH: 853
;; TYPE: PRT
;; ORGANISM: Escherichia coli
US-09-815-242-10272

Query Match 84.4%; Score 27; DB 10; Length 853;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALEKOL 7
DB 522 ALEKOL 527

RESULT 10
US-09-815-242-13869

;; Sequence 13869, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.

;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 13869
;; LENGTH: 855
;; TYPE: PRT
;; ORGANISM: Salmonella typhi
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (1)...(855)
;; OTHER INFORMATION: Xaa - Any Amino Acid
US-09-815-242-13869

Query Match 84.4%; Score 27; DB 10; Length 855;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALEKOL 7
DB 524 ALEKOL 529

RESULT 11
US-09-815-242-11081

;; Sequence 11081, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zyskind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; PRIOR FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11081
LENGTH: 861
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11081

Query Match 84.4%; Score 27; DB 10; Length 861;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALEKOL 7
DB 524 ALEKOL 529

RESULT 12
US-09-756-983-19
Sequence 19, Application US/09756983
Patent No. US20020122818A1
GENERAL INFORMATION:
APPLICANT: Albani, Salvatore
TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,
TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF
FILE REFERENCE: 246/285-CIP
CURRENT APPLICATION NUMBER: US/09/756,983
CURRENT FILING DATE: 2001-01-09
PRIOR APPLICATION NUMBER: 60/105,018
PRIOR FILING DATE: 1998-10-20
PRIOR APPLICATION NUMBER: 09/421,506
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: PCT/US99/2466
PRIOR FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 24
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19
LENGTH: 31
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptides
US-09-756-983-19

Query Match 81.2%; Score 26; DB 10; Length 31;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 23 QALEKEL 29

RESULT 13
US-09-766-378A-33
Sequence 33, Application US/09766378A
Patent No. US20020091079A1
GENERAL INFORMATION:
APPLICANT: Rhode, Peter R.
Acevedo, Jorge
Burkhardt, Martin
Jiao, Jin-an
Wong, Hing C.
TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND
METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP

STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: usa
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/766,378A
FILING DATE: 19-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/960,190
FILING DATE: 29-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Coiless, Peter F.
REGISTRATION NUMBER: 33, 860
REFERENCE/DOCKET NUMBER: 48002-DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-09-766-378A-33

Query Match 81.2%; Score 26; DB 10; Length 45;
Best Local Similarity 71.4%; Pred. No. 23;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 37 QALEKEL 43

RESULT 14
US-09-864-761-33411
Sequence 33411, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33411
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006440.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL - 0.67
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL - 1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL - 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL - 0.69
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL - 0.82
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL - 2.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL - 0.79
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL - 0.69
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL - 1.5
; OTHER INFORMATION: EST_HUMAN HIT: BE082334.1, EVALU8.00e-30
; OTHER INFORMATION: SWISSPROT HIT: P18175, EVALU8.1.10e-02
; US-09-864-761-33411

Query Match      81.2%  Score 26; DB 10; Length 67;
Best Local Similarity 85.7%  Pred. No. 34;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 EALEKOL 7
Db      35 ENLEKOL 41

RESULT 15
US-09-764-864-1269
; Sequence 1269, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1269
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (94)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
```

```

; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-864-1269

Query Match      81.2%  Score 26; DB 10; Length 122;
Best Local Similarity 85.7%  Pred. No. 63;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 EALEKOL 7
Db      28 ENLEKOL 34

Search completed: November 13, 2002, 13:40:18
Job time : 5.90881 secs
```


GENERAL INFORMATION:
APPLICANT: Benjamin, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Malne Edwards, Jean Baptiste
APPLICANT: Robert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81.052.REG
CURRENT APPLICATION NUMBER: US/09/834,366
CURRENT FILING DATE: 2001-04-13
PRIOR APPLICATION NUMBER: US 60/197,873
PRIORITY FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent.pm
SEQ ID NO 15179
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 21
OTHER INFORMATION: Xaa - Gly,Arg
US-09-834-366-15179

Query Match 100.0%; Score 32; DB 22; Length 141;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
DB 49 EALEKOL 55

RESULT 3
US-60-197-873-15179
Sequence 15179, Application US/60197873
GENERAL INFORMATION:
APPLICANT: Benjamin, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Malne Edwards, Jean Baptiste
APPLICANT: Robert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81.051.PRO
CURRENT APPLICATION NUMBER: US/60/197,873
CURRENT FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent.pm
SEQ ID NO 15179
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 21
OTHER INFORMATION: Xaa - Gly,Arg
US-60-197-873-15179

Query Match 100.0%; Score 32; DB 27; Length 141;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
DB 49 EALEKOL 55

RESULT 4
US-09-791-537-14762
Sequence 14762, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
FILE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 14762
LENGTH: 282
TYPE: PRT
ORGANISM: Mus musculus
US-09-791-537-14762

Query Match 100.0%; Score 32; DB 21; Length 282;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
DB 87 EALEKOL 93

RESULT 5
US-09-791-537-92908
Sequence 92908, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 92908
LENGTH: 542
TYPE: PRT
ORGANISM: Mus musculus
US-09-791-537-92908

Query Match 100.0%; Score 32; DB 21; Length 542;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
DB 327 EALEKOL 333

RESULT 6
US-09-791-537-125497
Sequence 125497, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 125497
LENGTH: 542
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-791-537-125497

Query Match 100.0%; Score 32; DB 21; Length 542;

Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
Db 328 EALEKOL 334

RESULT 7
US-10-205-331-57
Sequence 57, Application US/10205331
GENERAL INFORMATION:
APPLICANT: Warner-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Alister
APPLICANT: Brooksbank, Robert
APPLICANT: Plincock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WI-A-018199
CURRENT APPLICATION NUMBER: US/10/205,331
CURRENT FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: GB 0118354.0
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 117
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 57
LENGTH: 542
TYPE: PRT
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: NF-L
US-10-205-331-57

Query Match 100.0%; Score 32; DB 26; Length 542;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
Db 328 EALEKOL 334

RESULT 8
US-09-611-526-3132
Sequence 3132, Application US/09611526
GENERAL INFORMATION:
APPLICANT: OFA, TOSHIO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: ISOGAI, TAKAO
APPLICANT: HAYASHI, KOJI
APPLICANT: ISHII, SHIZUKO
APPLICANT: KAWAI, YURI
APPLICANT: WAKAMATSU, AI
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: NAGAI, KEIICHI
APPLICANT: KOJIMA, SHINICHI
APPLICANT: OTSUKI, TETSUJI
APPLICANT: KOGA, HISASHI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH CDNAS
TITLE OF INVENTION: AND THEIR USES
FILE REFERENCE: 08335/0122
CURRENT APPLICATION NUMBER: US/09/611,526
CURRENT FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: JP 1999-194486
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: JP 2000-118774
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183765
PRIOR FILING DATE: 2000-05-02
NUMBER OF SEQ ID NOS: 4484
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3132
LENGTH: 543

TYPE: PRT
ORGANISM: Homo sapiens
US-09-611-526-3132

Query Match 100.0%; Score 32; DB 20; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
Db 327 EALEKOL 333

RESULT 9
US-09-791-537-34475
Sequence 34475, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Danzer, Joseph
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 34475
LENGTH: 543
TYPE: PRT
ORGANISM: Mus musculus
US-09-791-537-34475

Query Match 100.0%; Score 32; DB 21; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
Db 328 EALEKOL 334

RESULT 10
US-09-791-537-56903
Sequence 56903, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Danzer, Joseph
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 56903
LENGTH: 543
TYPE: PRT
ORGANISM: Mus musculus
US-09-791-537-56903

Query Match 100.0%; Score 32; DB 21; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
Db 328 EALEKOL 334

RESULT 11
US-09-791-537-86802


```
; Sequence 86802, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 86802
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-86802
```

```
Query Match          100.0%; Score 32; DB 21; Length 544;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 EALEKOL 7
Db 328 EALEKOL 334
```

```
RESULT 12
US-09-791-537-125495
; Sequence 125495, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 125495
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-791-537-125495
```

```
Query Match          100.0%; Score 32; DB 21; Length 548;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 EALEKOL 7
Db 326 EALEKOL 332
```

```
RESULT 13
US-09-791-537-94465
; Sequence 94465, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 94465
; LENGTH: 554
; TYPE: PRT
```

```
; ORGANISM: Bos taurus
US-09-791-537-94465
```

```
Query Match          100.0%; Score 32; DB 21; Length 554;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 EALEKOL 7
Db 327 EALEKOL 333
```

```
RESULT 14
US-09-791-537-18635
; Sequence 18635, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18635
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-791-537-18635
```

```
Query Match          100.0%; Score 32; DB 21; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 EALEKOL 7
Db 328 EALEKOL 334
```

```
RESULT 15
US-09-791-537-30712
; Sequence 30712, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY )
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 30712
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Coturnix coturnix
US-09-791-537-30712
```

```
Query Match          100.0%; Score 32; DB 21; Length 556;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 EALEKOL 7
Db 329 EALEKOL 335
```

```
Search completed: November 13, 2002, 13:39:16
Job time : 141.447 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:17:03 ; Search time 2.82979 Seconds
(Without alignments)
29.431 Million cell updates/sec

Title: US-09-856-086-3

Perfect score: 32

Sequence: 1 EALEKQL 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 41632 seqs, 11897606 residues

Total number of hits satisfying chosen parameters: 41632

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/pdata/1/paa/pct_NEW_COMB.pep:*
2: /cgn2_6/pdata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/pdata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/pdata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/pdata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/pdata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/pdata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	100.0	7	US-09-856-086-3	Sequence 3, Appl 1
2	32	100.0	103	US-09-513-999C-6211	Sequence 6211, Ap
3	30	93.8	229	US-10-092-411A-5029	Sequence 5029, Ap
4	28	87.5	29	US-10-188-444-8	Sequence 8, Appl 1
5	28	87.5	29	US-10-103-597A-8	Sequence 8, Appl 1
6	27	84.4	159	US-10-125-923A-540	Sequence 540, App
7	27	84.4	240	US-10-120-801-69	Sequence 69, Appl 1
8	27	84.4	436	US-10-120-801-67	Sequence 68, Appl 1
9	27	84.4	1245	US-10-120-801-16	Sequence 67, Appl 1
10	27	84.4	1259	US-09-911-261A-19	Sequence 19, Appl 1
11	26	81.2	30	US-10-264-237-1810	Sequence 1810, Ap
12	26	81.2	161	US-10-092-411A-3338	Sequence 3338, Ap
13	26	81.2	535	US-10-092-411A-5144	Sequence 5144, Ap
14	26	81.2	652	US-10-113-709A-4	Sequence 4, Appl 1
15	25	78.1	14	US-10-271-078-28	Sequence 28, Appl 1
16	25	78.1	14	US-10-188-444-9	Sequence 9, Appl 1
17	25	78.1	29	US-10-103-597A-9	Sequence 9, Appl 1
18	25	78.1	90	US-10-092-411A-4461	Sequence 4461, Ap
19	25	78.1	150	US-10-092-411A-5144	Sequence 5144, Ap
20	25	78.1	307	US-10-113-709A-4	Sequence 4, Appl 1
21	25	78.1	393	US-10-137-953-11	Sequence 11, Appl 1
22	25	78.1	451	US-10-092-411A-4461	Sequence 4461, Ap
23	25	78.1	517	US-10-271-078-4	Sequence 4, Appl 1
24	25	78.1	591	US-10-271-078-2	Sequence 2, Appl 1
25	25	78.1	1689	US-10-080-943-2	Sequence 2, Appl 1
26	25	78.1	1698	US-10-080-943-4	Sequence 4, Appl 1

27	24	75.0	62	US-09-513-999C-6203	Sequence 6203, Ap
28	24	75.0	68	US-09-513-999C-7414	Sequence 7414, Ap
29	24	75.0	85	US-09-699-266A-4	Sequence 4, Appl 1
30	24	75.0	110	US-09-513-999C-5894	Sequence 5894, Ap
31	24	75.0	150	US-10-092-411A-5160	Sequence 5160, Ap
32	24	75.0	195	PCT-US02-32850-30	Sequence 30, Appl 1
33	24	75.0	195	US-10-264-237-2600	Sequence 2600, Ap
34	24	75.0	204	US-10-092-411A-4517	Sequence 4517, Ap
35	24	75.0	319	US-10-120-801-28	Sequence 28, Appl 1
36	24	75.0	321	US-10-120-801-92	Sequence 92, Appl 1
37	24	75.0	322	US-10-120-801-93	Sequence 93, Appl 1
38	24	75.0	347	US-10-092-411A-5508	Sequence 5508, Ap
39	24	75.0	387	US-10-240-842-4	Sequence 4, Appl 1
40	24	75.0	393	US-10-264-237-2785	Sequence 2785, Ap
41	24	75.0	408	US-10-092-411A-4130	Sequence 4130, Ap
42	24	75.0	478	US-10-154-678-108	Sequence 108, App
43	24	75.0	490	US-10-085-198-146	Sequence 146, App
44	24	75.0	490	US-10-085-198-148	Sequence 148, App
45	24	75.0	495	US-10-092-411A-4273	Sequence 4273, Ap

ALIGNMENTS

RESULT 1
US-09-856-086-3
Sequence 3, Application US/09856086
GENERAL INFORMATION:
APPLICANT: EBRINGER, ALAN
TITLE OF INVENTION: DIAGNOSIS OF DEMENTIATING OR SPONGIFORM DISEASE
FILE REFERENCE: 78104.040
CURRENT APPLICATION NUMBER: US/09/856,086
CURRENT FILING DATE: 2001-07-30
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 7
TYPE: PRT
ORGANISM: Homo sapiens, Bos
US-09-856-086-3

Query Match 100.0%; Score 32; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 3.8e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKQL 7
DB 1 EALEKQL 7
|||||
US-09-513-999C-6211
Sequence 6211, Application US/09513999C
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duchet, A.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
FILE REFERENCE: 59,US2,REG
CURRENT APPLICATION NUMBER: US/09/513,999C
PRIOR APPLICATION DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 6211
LENGTH: 103
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 90
OTHER INFORMATION: Xaa-Leu or Val

US-09-513-999C-6211

Query Match 100.0%; Score 32; DB 5; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
|||||
DB 81 EALEKOL 87

RESULT 3

US-10-092-411A-5029
; Sequence 5029, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 5029
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-5029

Query Match 93.8%; Score 30; DB 6; Length 229;
Best Local Similarity 85.7%; Pred. No. 9.3;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
|||||
DB 105 EALEKOL 111

RESULT 4

US-10-188-444-8
; Sequence 8, Application US/10188444
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142 (CIP)
; CURRENT APPLICATION NUMBER: US/10/188,444
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 9922352.1
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: amino acid sequence of leucine zippers
US-10-188-444-8

Query Match 87.5%; Score 28; DB 6; Length 29;
Best Local Similarity 85.7%; Pred. No. 3;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
|||||
DB 3 EALEKOL 9

RESULT 5
US-10-103-597A-8
; Sequence 8, Application US/10103597A
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142
; CURRENT APPLICATION NUMBER: US/10/103,597A
; PRIOR FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 9922352.1
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: amino acid sequence of leucine zippers
US-10-103-597A-8

Query Match 87.5%; Score 28; DB 6; Length 29;
Best Local Similarity 85.7%; Pred. No. 3;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
|||||
DB 3 EALEKOL 9

RESULT 6
US-10-125-923A-540
; Sequence 540, Application US/10125923A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C79
; CURRENT APPLICATION NUMBER: US/10/125,923A
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544

;; PRIOR FILING DATE: 1997-10-28
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 540
;; LENGTH: 159
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-125-923A-540

Query Match 84.4%; Score 27; DB 6; Length 159;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EALEKOL 7
DB 25 ESEKOL 31

RESULT 7
US-10-120-801-69
; Sequence 69, Application US/10120801
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyanekar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 69
; LENGTH: 240
; TYPE: PRT
; ORGANISM: human
US-10-120-801-69

Query Match 84.4%; Score 27; DB 6; Length 240;
Best Local Similarity 71.4%; Pred. No. 40;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
DB 63 ESEKOL 69

RESULT 8
US-10-120-801-68

; Sequence 68, Application US/10120801
; GENERAL INFORMATION:

; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyanekar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 68
; LENGTH: 436
; TYPE: PRT
; ORGANISM: human
US-10-120-801-68

Query Match 84.4%; Score 27; DB 6; Length 436;
Best Local Similarity 71.4%; Pred. No. 74;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
DB 259 ESEKOL 265

RESULT 9
US-10-120-801-67
; Sequence 67, Application US/10120801
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyanekar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11

PRIOR APPLICATION NUMBER: 60/285748
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 60/286068
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 60/286292
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/288334
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/291241
PRIOR FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/322284
PRIOR FILING DATE: 2001-09-14
PRIOR APPLICATION NUMBER: 60/285609
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 155
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 67
LENGTH: 1245
TYPE: PRT
ORGANISM: human
US-10-120-801-67

Query Match
Best Local Similarity 84.4%; Score 27; DB 6; Length 1245;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 1068 ESLEKOL 1074

RESULT 10
US-10-120-801-16
Sequence 16, Application US/10120801
GENERAL INFORMATION:
APPLICANT: Pena, Carol
APPLICANT: Guo, Xiaojia
APPLICANT: Shinkets, Richard
APPLICANT: Padigaru, Muralidhara
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytek, Kimberly
APPLICANT: Mehraban, Foad
APPLICANT: Topper, James N.
APPLICANT: Malvankar, Uriel
APPLICANT: Wasserman, Scott
APPLICANT: Edinger, Shlomit
APPLICANT: Smithson, Glenda
APPLICANT: Gunther, Erik
APPLICANT: Komuves, Laszlo
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-340
CURRENT APPLICATION NUMBER: US/10/120,801
CURRENT FILING DATE: 2002-04-11
PRIOR APPLICATION NUMBER: 60/285748
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 60/286068
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 60/286292
PRIOR FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/288334
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/291241
PRIOR FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 60/322284
PRIOR FILING DATE: 2001-09-14
PRIOR APPLICATION NUMBER: 60/285609
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 155
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 1259
TYPE: PRT
ORGANISM: Homo sapiens

US-10-120-801-16

Query Match
Best Local Similarity 84.4%; Score 27; DB 6; Length 1259;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 1082 ESLEKOL 1088

RESULT 11
US-09-911-261A-19
Sequence 19, Application US/09911261A
GENERAL INFORMATION:
APPLICANT: Sera, Takashi
TITLE OF INVENTION: Zinc Finger Domain Recognition Code and Uses Thereof
FILE REFERENCE: 109845.135
CURRENT APPLICATION NUMBER: US/09/911,261A
CURRENT FILING DATE: 2001-07-23
PRIOR APPLICATION NUMBER: US 60/220,060
PRIOR FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 69
SOFTWARE: PatentIn version 3.0
SEQ ID NO 19
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Acid dimerization peptide.
US-09-911-261A-19

Query Match
Best Local Similarity 81.2%; Score 26; DB 5; Length 30;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 22 QALEKOL 28

RESULT 12
US-10-264-237-1810
Sequence 1810, Application US/10264237
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA13191
CURRENT APPLICATION NUMBER: US/10/264,237
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/16450
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: US 60/205,515
PRIOR FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 2876
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 1810
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-237-1810

Query Match
Best Local Similarity 81.2%; Score 26; DB 6; Length 161;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 153 EAVEROL 159

RESULT 13
US-10-092-411A-3338

```

; Sequence 3338, Application US/10092411A
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: 032796-101
; CURRENT APPLICATION NUMBER: US/10/092,411A
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5676
; SEQ ID NO 3338
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-092-411A-3338

```

```

Query Match          81.2%; Score 26; DB 6; Length 535;
Best Local Similarity 71.4%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 EALEKOL 7
      |||||:|
DB      185 EALEREL 191

```

```

RESULT 14
PCT-US02-32851-17
; Sequence 17, Application PC/TUS0232851
; GENERAL INFORMATION:

```

```

; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BECHTA, Shanya D.
; APPLICANT: BHATIA, Umesh
; APPLICANT: BLAKE, Julie J.
; APPLICANT: BOROMSKY, Mark L.
; APPLICANT: BURRILL, John D.
; APPLICANT: CHANG, Hsin-Ru
; APPLICANT: CHAMLA, Narinder K.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: EMERLING, Brooke M.
; APPLICANT: FORSYTHE, Ian J.
; APPLICANT: GORVAD, Ann E.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: HO, Anne
; APPLICANT: ISON, Craig H.
; APPLICANT: KABLE, Amy E.
; APPLICANT: KHARE, Reema
; APPLICANT: LAL, Preeti G.
; APPLICANT: LEE, Sally
; APPLICANT: LEE, Ernestine A.
; APPLICANT: LEE, Soo Yeun
; APPLICANT: LEHR-MASON, Patricia M.
; APPLICANT: LI, Joana X.
; APPLICANT: LINDQUIST, Erika A.
; APPLICANT: LUD, Wen
; APPLICANT: MARQUIS, Joseph P.
; APPLICANT: RAMAKRISHNAN, Jayalakshmi
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: SPRAGUE, William W.
; APPLICANT: SWARNAKAR, Anita
; APPLICANT: TANG, Y. Tom
; APPLICANT: WARREN, Bridget A.
; APPLICANT: YANG, Junming
; APPLICANT: YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; APPLICANT: ZHENG, Wenjin
; TITLE OF INVENTION: STRUCTURAL AND CYTOSKELETON-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1223 PCT

```

```

; CURRENT APPLICATION NUMBER: PCT/US02/32851
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 60/328,931
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/360,681
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 60/343,896
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 60/346,308
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/332,385
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/340,776
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/347,703
; PRIOR FILING DATE: 2002-01-11
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7502375CD1
PCT-US02-32851-17

```

```

Query Match          81.2%; Score 26; DB 1; Length 652;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 EALEKOL 7
      |||||:|
DB      353 EALEKEY 359

```

```

RESULT 15
US-10-271-078-28
; Sequence 28, Application US/10271078
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN LEUCINE-RICH REPEAT DOMAIN CONTAINING PROTEIN, HLL
; FILE REFERENCE: D0157 NP
; CURRENT APPLICATION NUMBER: US/10/271,078
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: U.S. 60/328,478
; PRIOR FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-271-078-28

```

```

Query Match          78.1%; Score 25; DB 6; Length 14;
Best Local Similarity 71.4%; Pred. No. 6.1;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 EALEKOL 7
      |||||:|
DB      2 ESEKRL 8

```

```

Search completed: November 13, 2002, 13:39:38
Job time : 2.82979 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 ; Search time 10.8723 seconds
(without alignments)
61.895 Million cell updates/sec

Title: US-09-856-086-3

Perfect score: 32

Sequence: 1 EALEKOL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	100.0	287	2 A21762	neurofilament trip
2	32	100.0	543	1 OFMSL	neurofilament trip
3	32	100.0	544	2 S07144	neurofilament trip
4	32	100.0	548	1 OFPGL	neurofilament trip
5	32	100.0	554	2 JN0094	neurofilament prot
6	32	100.0	556	2 B46024	neurofilament-L su
7	32	100.0	3660	1 S02041	dystrophin, muscle
8	30	93.8	384	2 T44652	UDP-N-acetylglucos
9	30	93.8	451	2 F83747	denosylmethionine
10	29	90.6	166	2 D37844	base 19.5K protein
11	29	90.6	249	2 T04307	M79 protein - rice
12	29	90.6	249	2 T04335	MADS box protein -
13	29	90.6	253	2 H70188	conserved hypobeth
14	29	90.6	259	2 T04169	MADS box protein -
15	29	90.6	284	2 T17222	hypothetical prote
16	29	90.6	337	2 C71003	hypothetical prote
17	29	90.6	382	2 B40838	GA-binding protein
18	29	90.6	395	2 I38741	nuclear respirator
19	29	90.6	494	2 JN0065	anuranlate synth
20	29	90.6	551	2 T40767	hypothetical prote
21	29	90.6	727	2 T00383	hypothetical prote
22	29	90.6	805	2 T41810	KINA0632 protein -
23	29	90.6	808	2 C72858	ACMPV orf66 - Bom
24	29	90.6	1078	2 T18352	Acortf-66 protein -
25	29	90.6	1079	2 T18356	protein p120 - Myc
26	29	90.6	1232	2 B39432	membrane protein p
27	29	90.6	1520	1 TVEFA	ATP-dependent deox
28	29	90.6	1922	2 T00637	protein-tyrosine k
29	28	87.5	51	2 A70229	hypothetical prote
30	28	87.5	51	2 A70229	hypothetical prote

30	28	87.5	72	2 B70256	hypothetical prote
31	28	87.5	116	2 A71565	hypothetical prote
32	28	87.5	184	2 F84584	hypothetical prote
33	28	87.5	219	2 H97110	phosphate uptake r
34	28	87.5	240	2 AD0534	conserved hypobeth
35	28	87.5	250	2 AG2551	chromosome partitl
36	28	87.5	280	2 T47660	NAM (no apical mer
37	28	87.5	308	2 H81345	hypothetical prote
38	28	87.5	330	2 A12820	two component sens
39	28	87.5	336	2 AH3264	cysteine synthase
40	28	87.5	339	2 S76619	hypothetical prote
41	28	87.5	342	2 A97599	sensor histidine k
42	28	87.5	348	2 G85661	hypothetical prote
43	28	87.5	372	2 G64064	outer membrane int
44	28	87.5	492	2 AH1030	probable exported
45	28	87.5	501	2 T00213	type II secretion

ALIGNMENTS

RESULT 1
A21762
neurofilament triplet L protein - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 13-Aug-1999
C:Accession: A21762
R;Jullien, J.P.; Ramchandran, K.; Grosvel, F.
Biochim. Biophys. Acta 825, 398-404, 1985
A:Title: Cloning of a cDNA encoding the smallest neurofilament protein from the rat.
A:Reference number: A21762; MUID:85252830; PMID:3925999
A:Accession: A21762
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-287 <LE2>
A:Cross-references: GB:M25638; NID:9205683; PIDN:AAA1694.1; PID:9205684
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil

Query Match 100.0%; Score 32; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 131 EALEKOL 137

RESULT 2
OFMSL
neurofilament triplet L protein - mouse
N:Alternate names: 68k neurofilament protein; NF-L(low) protein; type IV IF protein
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1988 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999
C:Accession: A25227; A26562; A43772; A41012; I55316
R;Lewis, S.A.; Cowan, N.J.
Mol. Cell. Biol. 6, 1529-1534, 1986
A:Title: Anomalous placement of introns in a member of the intermediate filament mult
A:Reference number: A25227; MUID:87064433; PMID:3785173
A:Accession: A25227
A:Molecule type: DNA
A:Residues: 1-543 <LE2>
A:Cross-references: GB:M3016; NID:9200023; PIDN:AAA39610.1; PID:9387492
A>Note: the authors translated the codon GGC for residue 5 as Ala, ACA for residue 88
1 as Glu
R;Lewis, S.A.; Cowan, N.J.
J. Cell Biol. 100, 843-850, 1985
A:Title: Genetics, evolution, and expression of the 68,000-mol-wt neurofilament prote
A:Reference number: A26562; MUID:8513334; PMID:3919033
A:Accession: A26562
A:Molecule type: mRNA
A:Residues: 242-543 <LE2>
A:Cross-references: GB:X02165

A:Experimental source: brain
 A:Jullien, J.P.; Meyer, D.; Flavell, D.; Hurst, J.; Grosveld, F.
 Brain Res. Mol. Brain Res. 1, 243-250, 1986
 A:Title: Cloning and developmental expression of the murine neurofilament gene family.
 A:Reference number: A43772
 A:Accession: A43772
 A:Molecule type: mRNA
 A:Residues: 1-5, 'Y', '7-8, 'Y', '10-64, 'W', '66-72, 'L', '74-98, 'D', '100-194, 'R', '196-202, '204-239, 'Y'
 A:Cross-references: GB:M20480; NID:g200037; PIDN:AAA39814.1; PID:g200038
 A:Note: the authors translated the codon CGC for residue 195 as Ala
 R:Shay, R.K.; Nixon, R.A.
 J. Biol. Chem. 266, 18861-18867, 1991
 A:Title: Identification of Ser-55 as a major protein kinase A phosphorylation site on th
 A:Reference number: A41012; MUID:92011653; PMID:1717455
 A:Accession: A41012
 A:Molecule type: protein
 A:Residues: 52-57 <STH>
 R:Nakamura, K.; Ikenaka, K.; Wada, K.; Tamura, T.
 J. Biol. Chem. 265, 19786-19791, 1990
 A:Title: Structure of the 68-kDa neurofilament gene and regulation of its expression.
 A:Reference number: I55316; MUID:91060592; PMID:2246261
 A:Accession: I55316
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-5, 'Y', '7-8, 'Y', '10-28 <RES>
 A:Cross-references: GB:M55423; NID:g200027; PIDN:AAA39812.1; PID:g554245
 A:Comment: This is the most abundant of the three neurofilament proteins and, as the oth
 C:Genetics:
 A:Introns: 349/3; 391/2; 498/1
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:2-72/Domain: head <HED>
 F:94-135/Domain: coil 1a, alpha-helical rod #status predicted <R1a>
 F:126-138/Region: linker 1
 F:139-334/Domain: coil 1b, alpha-helical rod #status predicted <R1b>
 F:235-256/Region: linker 12
 F:257-372/Domain: coil 2a, alpha-helical rod #status predicted <R2a>
 F:273-281/Region: linker 2
 F:282-401/Domain: coil 2b, alpha-helical rod #status predicted <R2b>
 F:404-543/Domain: tail <TAI>
 F:404-444/Region: tail subdomain a
 F:445-543/Region: tail subdomain b

Query Match 100.0%; Score 32; DB 1; Length 543;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EALEKOL 7
 |||||
 Db 328 EALEKOL 334

RESULT 3
 S07144
 neurofilament triplet L protein - human
 N:Alternate names: neurofilament light polypeptide (68K)
 N:Contents: Glu-50 brain peptide
 C:Species: Homo sapiens (man)
 C:Date: 29-Jan-1993 #sequence; revision 29-Jan-1993 #text_change 21-Jul-2000
 C:Accession: S07144; I52832; A60703
 R:Jullien, J.P.; Grosveld, F.; Yazdankhah, K.; Flavell, D.; Meijer, D.; Mushynski, W.
 Blochim. Biophys. Acta 909, 10-20, 1987
 A:Title: The structure of a human neurofilament gene (NF-L): a unique exon-intron organ
 A:Reference number: S07144; MUID:87214213; PMID:3034332
 A:Accession: S07144
 A:Molecule type: DNA
 A:Residues: 1-544 <TUL>
 A:Cross-references: EMBL:X05608; NID:g1495072; PIDN:CAA29097.1; PID:g1279504
 A:Note: the authors translated the codon ATG for residue 366 as Asn
 R:Pospelov, V.A.; Pospelova, T.V.; Jullien, J.P.
 Cell Growth Differ. 5, 187-196, 1994
 A:Title: AP-1 and Krox-24 transcription factors activate the neurofilament light gene pr
 A:Reference number: I52832; MUID:94235564; PMID:8180132

A:Accession: I52832
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-10 <POS>
 A:Cross-references: GB:S70309; NID:g547176; PIDN:AD14057.1; PID:g4261757
 R:Nomata, Y.; Matsubara, T.; Wada, H.
 J. Biochem. 93, 825-831, 1983
 A:Title: Highly acidic proteins from human brain: purification and properties of Gl
 A:Reference number: A60703; MUID:83265667; PMID:6135695
 A:Accession: A60703
 A:Molecule type: protein
 A:Residues: 469-472, 'D', '474 <NOM>
 A:Experimental source: Glu-50 brain peptide
 A:Note: this acidic protein is named for its greater than fifty per cent glutamic a
 C:Genetics:
 A:Gene: GDB:NEFL; NFL
 A:Cross-references: GDB:120227; OMTN:162280
 A:Map position: 8p21-8p21
 A:Introns: 349/3; 391/2; 498/1
 C:Superfamily: cytoskeletal keratin
 C:Keywords: brain; coiled coil; intermediate filament
 F:469-544/Product: Glu-50 peptide #status predicted <ES0>

Query Match 100.0%; Score 32; DB 2; Length 544;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 EALEKOL 7
 |||||
 Db 328 EALEKOL 334

RESULT 4
 OEPL
 neurofilament triplet L protein - pig
 N:Alternate names: 68K neurofilament protein
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 15-Nov-1984 #sequence; revision 28-May-1986 #text_change 10-Oct-1997
 C:Accession: A91337; A90973; A34569; A02963
 R:Geisler, N.; Plessmann, U.; Weber, K.
 FEBS Lett. 182, 475-478, 1985
 A:Title: The complete amino acid sequence of the major mammalian neurofilament prot
 A:Reference number: A91337; MUID:85154583; PMID:3920075
 A:Accession: A91337
 A:Molecule type: protein
 A:Residues: 1-547 <GEI>
 R:Geisler, N.; Kaufmann, E.; Fischer, S.; Plessmann, U.; Weber, K.
 EMBO J. 2, 1295-1302, 1983
 A:Title: Neurofilament architecture combines structural principles of intermediate:
 A:Reference number: A90973
 A:Accession: A90973
 A:Molecule type: protein
 A:Residues: 1-82; 278-548 <GE2>
 A:Note: residue 372 is either lysine or arginine
 R:Gonda, Y.; Nishizawa, K.; Ando, S.; Kitamura, S.; Minoura, Y.; Nishl, Y.; Inagaki.
 Blochim. Biophys. Res. Commun. 167, 1316-1325, 1990
 A:Title: Involvement of protein kinase C in the regulation of assembly-disassembly
 A:Reference number: A34569; MUID:90211318; PMID:2108674
 A:Accession: A34569
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 9-14; 23-29; 30-53 <GON>
 C:Comment: Mammalian neurofilaments usually contain three polypeptides, L, M, and H
 ke all other intermediate filament proteins: a conserved alpha-helical region, whos
 C:Comment: The amino-terminal headpiece is basic with a high content of hydroxylamin
 al beta turns; domain b is acidic and rich in glutamic acid and lysine residues.
 C:Comment: The extra mass and high charge density that distinguish the neurofilament
 charged scaffolding structure suitable for interaction with other neuronal componen
 C:Comment: The boundaries of the domains for interaction with other neuronal componen
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:1-70/Domain: head <HED>

F:92-123/Domain: coil 1a, alpha-helical rod <R1A>
 F:137-232/Domain: coil 1b, alpha-helical rod <R1B>
 F:257-399/Domain: coil 2, alpha-helical rod <R2>
 F:402-548/Domain: tail <TA1>
 F:443-548/Region: tail subdomain a
 F:443-548/Region: tail subdomain b

Query Match 100.0%; Score 32; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
 |||||
 DB 326 EALEKOL 332

RESULT 5
 JM0094
 neurofilament protein-L - bovine
 N:Alternate names: NF-L
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 07-May-1999
 C:Accession: JM0094
 R:Hashimoto, R.; Nakamura, Y.; Goto, H.; Wada, Y.; Sakoda, S.; Kalbuch, K.; Inagaki, M.
 Biochem. Biophys. Res. Commun. 245, 407-411, 1998
 A:Title: Domain- and site-specific phosphorylation of bovine NF-L by Rho-associated kinase
 A:Reference number: JM0094; PMID:98238650; PMID:9571164
 A:Accession: JM0094
 A:Molecule type: protein
 A:Residues: 1-554 <RHS>
 C:Comment: Domain- and site-specific phosphorylation by Rho-kinase regulates the assembly
 C:Superfamily: cytoskeletal keratin

Query Match 100.0%; Score 32; DB 2; Length 554;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
 |||||
 DB 327 EALEKOL 333

RESULT 6
 B46024
 neurofilament-L subunit - quail
 C:Species: Coturnix coturnix (quail)
 C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
 C:Accession: B46024
 R:Ohara, O.; Gabarra, Y.; Miyake, T.; Teraoka, H.; Kitamura, T.
 J. Cell Biol. 121, 387-395, 1993
 A:Title: Neurofilament deficiency in quail caused by nonsense mutation in neurofilament-
 A:Reference number: A46024; PMID:9324534; PMID:8468353
 A:Accession: B46024
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-556 <DNA>
 A:Experimental source: subsp. japonica, TRP
 A:Note: sequence inconsistent with the nucleotide translation
 A:Note: sequence extracted from NCBI backbone (NCBIN:129455, NCBI:129482)
 C:Superfamily: cytoskeletal keratin

Query Match 100.0%; Score 32; DB 2; Length 556;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
 |||||
 DB 329 EALEKOL 335

RESULT 7
 S02041
 dystrophin, muscle - chicken

N:Alternate names: duchenne muscular dystrophy protein
 C:Species: Gallus gallus (chicken)
 C>Date: 07-Sep-1990 #sequence_revision 27-Jun-1994 #text_change 16-Jul-1999
 C:Accession: S02041; S02013; S71487
 R:Lemaire, C.; Heilig, R.; Mandel, J.L.
 Nucleic Acids Res. 16, 11815-11816, 1988
 A:Title: Nucleotide sequence of chicken dystrophin cDNA.
 A:Reference number: S02041; PMID:89098331; PMID:3062582
 A:Accession: S02041

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-3660 <LEM>

A:Cross-references: EMBL:X13369; NID:963369; PIDN:CAA31746.1; PID:963370
 A:Note: 1869-His, 1885-Arg, and sequences lacking 1171-Met were also found

R:Lemaire, C.; Heilig, R.; Mandel, J.L.
 EMBO J. 7, 4157-4162, 1988

A:Title: The chicken dystrophin cDNA: striking conservation of the C-terminal coding
 A:Reference number: S02013; PMID:89210800; PMID:3072195
 A:Accession: S02013

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-3573, 'HA', 3576-3660 <LEM2>

R:Heilig, R.; Lemaire, C.; Mandel, J.L.
 Nucleic Acids Res. 15, 9129-9142, 1987

A:Title: A 230kb cosmid walk in the Duchenne muscular dystrophy gene: detection of a
 A:Reference number: S09071; PMID:88067745; PMID:2825128
 A:Accession: S71487

A:Molecule type: DNA

A:Residues: 222-281 <HE1>

C:Comment: Dystrophin is proposed to play a role in anchoring the cytoskeleton to the
 C:Comment: Defects in dystrophin are responsible for the Duchenne/Becker muscular dys

C:Superfamily: dystrophin; alpha-actinin actin-binding domain homology; spectrin/dyst

C:Keywords: actin binding; calmodulin binding; cytoskeleton; leucine zipper; membrane

F:18-237/Domain: alpha-actinin actin-binding domain homology <ACT>

F:253-327/Region: hinge

F:340-449/Domain: spectrin/dystrophin repeat homology <SP01>

F:450-558/Domain: spectrin/dystrophin repeat homology <SP02>

F:560-669/Domain: spectrin/dystrophin repeat homology <SP03>

F:670-719/Region: hinge

F:720-830/Domain: spectrin/dystrophin repeat homology <SP04>

F:838-936/Domain: spectrin/dystrophin repeat homology <SP05>

F:940-1047/Domain: spectrin/dystrophin repeat homology <SP06>

F:1049-1156/Domain: spectrin/dystrophin repeat homology <SP07>

F:1158-1265/Domain: spectrin/dystrophin repeat homology <SP08>

F:1267-1369/Domain: spectrin/dystrophin repeat homology <SP09>

F:1374-1479/Domain: spectrin/dystrophin repeat homology <SP10>

F:1480-1570/Domain: spectrin/dystrophin repeat homology <SP11>

F:1572-1678/Domain: spectrin/dystrophin repeat homology <SP12>

F:1680-1784/Domain: spectrin/dystrophin repeat homology <SP13>

F:1787-1877/Domain: spectrin/dystrophin repeat homology <SP14>

F:1878-1984/Domain: spectrin/dystrophin repeat homology <SP15>

F:1986-2103/Domain: spectrin/dystrophin repeat homology <SP16>

F:2105-2211/Domain: spectrin/dystrophin repeat homology <SP17>

F:2213-2319/Domain: spectrin/dystrophin repeat homology <SP18>

F:2323-2419/Domain: spectrin/dystrophin repeat homology <SP19>

F:2420-2467/Region: hinge

F:2468-2574/Domain: spectrin/dystrophin repeat homology <SP20>

F:2576-2683/Domain: spectrin/dystrophin repeat homology <SP21>

F:2685-2799/Domain: spectrin/dystrophin repeat homology <SP22>

F:2801-2928/Domain: spectrin/dystrophin repeat homology <SP23>

F:2930-3037/Domain: spectrin/dystrophin repeat homology <SP24>

F:3038-3075/Region: hinge

F:3082-3089/Domain: WW repeat homology <WW1>

F:3079-3357/Region: cysteine-rich

F:3481-3502/Region: leucine zipper motif

F:3547-3568/Region: leucine zipper motif

Query Match 100.0%; Score 32; DB 1; Length 3660;
 Best Local Similarity 100.0%; Pred. No. 4,440+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
 |||||

Db 1084 EALEKOL 1090

RESULT 8

T44652

UDP-N-acetylglucosamine-2-epimerase cpso [Imported] - Streptococcus agalactiae

C:Species: Streptococcus agalactiae

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000

C:Accession: T44652

R:Chaffin, D.O.; Yim, H.H.; Beres, S.B.; Sweet, E.S.; Nittayaajarn, A.; Rubens, C.E.

submitted to the EMBL Data Library, June 1999

A:Reference number: 222821

A:Accession: T44652

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-384 <CHA>

A:Cross-references: EMBL:AF163833; PIDN:AD53075.1

A:Experimental source: strain COH1, serotype III

C:Genetics:

A:Gene: cpso

Query Match 93.8%; Score 30; DB 2; Length 384;

Best Local Similarity 85.7%; Pred. No. 1e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7

Db 190 EALEKOL 196

RESULT 9

F83747

denosylmethionine-8-amino-7-oxononanoate aminotransferase BH0782 [Imported] - Bacillus h

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 17-May-2002

C:Accession: F83747

R:Takami, H.; Nakasone, K.; Takaki, Y.; Mieno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: F83747

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-451 <STO>

A:Cross-references: GB:AP001509; GB:BA000004; NID:910173176; PIDN:BA04501.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH0782

C:Superfamily: beta-alanine-pyruvate transaminase

Query Match 93.8%; Score 30; DB 2; Length 451;

Best Local Similarity 85.7%; Pred. No. 1.2e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7

Db 68 EALEKOL 74

RESULT 10

D37844

bale 19.5K protein - Eubacterium sp.

C:Species: Eubacterium sp.

C:Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 22-Oct-1999

C:Accession: D37844

R:Malinow, D.H.; White, W.B.; Hylemon, P.B.

J. Bacteriol. 172, 7011-7019, 1990

A:Title: Cloning and sequencing of a bile acid-inducible operon from Eubacterium sp. str

A:Reference number: A37844; MUID:91072253; PMID:2254270

A:Accession: D37844

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-166 <MAL>

A:Cross-references: GB:U57489; GB:M36292; NID:g1381562; PIDN:AAC45413.1; PID:g13815

Query Match 90.6%; Score 29; DB 2; Length 166;

Best Local Similarity 85.7%; Pred. No. 69;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7

Db 8 EALEKOL 14

RESULT 11

T04307

M79 protein - rice

C:Species: Oryza sativa (rice)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000

C:Accession: T04307

R:Qu, L.J.

submitted to the EMBL Data Library, September 1997

A:Reference number: Z15271

A:Accession: T04307

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-249 <QUL>

A:Cross-references: EMBL:Y15008; PIDN:CAA75241.1

A:Experimental source: subsp. Japonica, cv. Zhonghua 8, flower

C:Genetics:

A:Gene: M79

C:Superfamily: transcription factor equa; serum response factor DNA-binding domain

F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match 90.6%; Score 29; DB 2; Length 249;

Best Local Similarity 85.7%; Pred. No. 1.1e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7

Db 128 ESKOL 134

RESULT 12

T04335

MADS box protein - rice

C:Species: Oryza sativa (rice)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000

C:Accession: T04335

R:Greco, R.; Stagi, L.; Colombo, L.; Angenent, G.C.; Sari-Gorla, M.; Pe, M.E.

Mol. Gen. Genet. 253, 615-623, 1997

A:Title: MADS box genes expressed in developing inflorescences of rice and sorghum.

A:Reference number: Z15292; MUID:97218034; PMID:9065695

A:Accession: T04335

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-249 <GRE>

A:Cross-references: EMBL:U31994; NID:g1905929; PIDN:AAB50180.1; PID:g1905930

C:Genetics:

A:Note: MADS45

C:Function:

A:Description: may be involved in floral development as a transcription factor

C:Superfamily: transcription factor equa; serum response factor DNA-binding domain

C:Keywords: DNA binding; transcription regulation

F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match 90.6%; Score 29; DB 2; Length 249;

Best Local Similarity 85.7%; Pred. No. 1.1e+02;

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7

Db 128 ESKOL 134

RESULT 13

H70188
 conserved hypothetical protein BB0713 - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C>Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
 C:Accession: H70188
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
 son, D.; Peterson, J.; Klevans, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
 Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
 A:Accession: H70188
 A:Reference number: A70100; MUID:98065943; PMID:9403685
 A:Molecule type: DNA
 A:Residues: 1-253 <KLE>
 A:Cross-references: GB:AE001171; GB:AE000783; NID:92688640; PIDN:AC67060.1; PID:9268864
 A:Experimental source: strain B31

Query Match 90.6%; Score 29; DB 2; Length 253;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
 |||||:
 DB 92 EALEKEL 98

RESULT 14

T04169
 MADS box protein - rice
 C:Species: Oryza sativa (rice)
 C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000
 C:Accession: T04169
 R:Kang, H.G.; Jang, S.; Chung, J.E.; Cho, Y.G.; An, G.
 Molecules and Cells 7, 559-566, 1997
 A>Title: Characterization of two rice MADS box genes that control flowering time.
 A:Reference number: Z15253; MUID:97479823; PMID:9339904
 A:Accession: T04169
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-259 <KAN>
 A:Cross-references: EMBL:U78891; NID:92286110; PID:92286111
 A:Experimental source: strain M201
 C:Genetics:
 A:Gene: MADS7
 C:Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
 F:13-67/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match 90.6%; Score 29; DB 2; Length 259;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
 |||||:
 DB 138 ESEKOL 144

RESULT 15

T17222
 hypothetical protein DKFZp564g2263.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T17222
 R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, September 1999
 A:Reference number: Z18727
 A:Accession: T17222
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-284 <DUE>
 A:Cross-references: EMBL:AL117418
 A:Experimental source: fetal brain; clone DKFZp564g2263

C:Genetics:
 A>Note: DKFZp564g2263.1

Query Match 90.6%; Score 29; DB 2; Length 284;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
 |||||:
 DB 271 QALEKOL 277

Search completed: November 13, 2002, 13:23:26
 Job time: 12.0152 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 5.80851 Seconds
(without alignments)
49.984 Million cell updates/sec

Title: US-09-856-086-3
Perfect score: 32
Sequence: 1 EALEKOL 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	100.0	541	1	NFL_RAT
2	32	100.0	542	1	NFL_MOUSE
3	32	100.0	543	1	NFL_HUMAN
4	32	100.0	548	1	NFL_PIG
5	32	100.0	554	1	NFL_BOVIN
6	32	100.0	555	1	NFL_CORTA
7	32	100.0	3660	1	DMD_CHICK
8	29	90.6	166	1	BALF_EBSP
9	29	90.6	354	1	REF1_THERM
10	29	90.6	382	1	GABR_MOUSE
11	29	90.6	383	1	GABR_HUMAN
12	29	90.6	414	1	OAT_PLAFD
13	29	90.6	484	1	TRPE_CLODM
14	29	90.6	808	1	Y066_NPVAC
15	29	90.6	1087	1	AKA8_RABIT
16	29	90.6	1232	1	ADDA_BACSU
17	29	90.6	1520	1	ABL_DROME
18	29	90.6	3911	1	AKA9_HUMAN
19	28	87.5	287	1	FTF_METBA
20	28	87.5	316	1	RPOA_MASVI
21	28	87.5	339	1	Y090_STYX
22	28	87.5	372	1	TOLA_HAELN
23	28	87.5	407	1	PSR3_HAELN
24	28	87.5	550	1	MANB_MYGE
25	28	87.5	575	1	ACER_YEAST
26	28	87.5	1012	1	PROL_MOUSE
27	28	87.5	1013	1	PROL_HUMAN
28	28	87.5	1013	1	PROL_RAT
29	28	87.5	1189	1	ALAE_ARATH
30	28	87.5	1213	1	RPOC_STRPY
31	27	84.4	133	1	TX12_HUMAN
32	27	84.4	123	1	TX12_MOUSE
33	27	84.4	186	1	PSAF_CTAFA

34	27	84.4	187	1	SIGW_BACSU	045585 bacillus su
35	27	84.4	192	1	LPCA_HELPJ	026221 helicobacte
36	27	84.4	192	1	LPCA_HELPJ	025528 helicobacte
37	27	84.4	207	1	HIS2_CAMJE	09pm71 campylobact
38	27	84.4	247	1	YQ97_YERPE	082d91 yersinia pe
39	27	84.4	289	1	ATPG_MYCGA	P33251 mycoplasma
40	27	84.4	340	1	Y269_MYCGE	049407 mycoplasma
41	27	84.4	366	1	YF72_HAELN	P46495 haemophilus
42	27	84.4	375	1	AGP2_BOVIN	077802 bos taurus
43	27	84.4	383	1	RNR_VIBRA	P40611 vibrio para
44	27	84.4	428	1	INVO_CERAB	P24705 cebus albid
45	27	84.4	493	1	INVO_SAGOE	P24712 sequinus oe

ALIGNMENTS

RESULT 1	ID	NFL_RAT	STANDARD	PRT	541 AA.
AC	P19527	063367			
DT	01-FEB-1991	(Rel. 17, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Neurofilament triplet L protein (68 kDa neurofilament protein)				
DE	(Neurofilament light polypeptide) (NF-L).				
GN	NFL OR NFL OR NF68.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90184052; PubMed=2516804;				
RA	Chin S.S., Liem R.K.H.,				
RT	"Expression of rat neurofilament proteins NF-L and NF-M in				
RT	transfected non-neuronal cells."				
RL	Eur. J. Cell Biol. 50:475-490(1989).				
RN	[2]				
RP	SEQUENCE OF 197-483 FROM N.A.				
RX	MEDLINE=85252830; PubMed=3925999;				
RA	Julien J.-P., Ramachandran K., Grosfeld F.;				
RT	"Cloning of a cDNA encoding the smallest neurofilament protein from				
RT	the rat."				
RL	Biochim. Biophys. Acta 825:398-404(1985).				
RN	[3]				
RP	SEQUENCE OF 1-10 FROM N.A.				
RX	MEDLINE=95264348; PubMed=7745611;				
RA	Reeben M., Neuman T., Palgi J., Palm K., Paalme V., Saarma M.;				
RT	"Characterization of the rat light neurofilament (NF-L) gene promoter				
RT	and identification of NGF and cAMP responsive regions."				
RL	J. Neurosci. Res. 40:177-188(1995).				
RN	[4]				
RP	CARBOHYDRATE-LINKAGE SITES.				
RX	MEDLINE=93346421; PubMed=8344946;				
RA	Dong D.L.-Y., Xu Z.-S., Chevalier M.R., Cotter R.J., Cleveland D.W.,				
RT	Hart G.W.;				
RT	"Glycosylation of mammalian neurofilaments. Localization of multiple				
RT	O-linked N-acetylglucosamine moieties on neurofilament polypeptides				
RT	L and M."				
RL	J. Biol. Chem. 268:16679-16687(1993).				
CC	-1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,				
CC	AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.				
CC	-1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH				
CC	THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT				
CC	PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM				
CC	A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH				
CC	OTHER NEURONAL COMPONENTS OR IONS.				
CC	-1- PTM: O-GLYCOSYLATED; CONTAINS THREE N-ACETYLGLUCOSAMINE SIDE				
CC	CHAINS.				
CC	-1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE				
CC	NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL				
CC	INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM				

CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF031880; AAB67069.1; -
 CC EMBL: M25638; AAA41694.1; -
 CC EMBL: X53981; CAA37931.1; -
 CC PIR: A21762; A21762.
 CC Glycositedb: P19527; -
 CC Interpro: IPR001664; IF.
 CC Pfam: PF00038; filament; 1.
 CC Prosite: PS00226; IF; 1.
 CC Intermediate filament; Coiled coil; Neurone; Glycoprotein.
 CC INT_MET 0 0
 CC FT DOMAIN 1 92 HEAD.
 CC FT DOMAIN 93 396 ROD.
 CC FT DOMAIN 397 541 TAIL.
 CC FT DOMAIN 93 124 COIL 1A.
 CC FT DOMAIN 125 137 LINKER 1.
 CC FT DOMAIN 138 233 COIL 1B.
 CC FT DOMAIN 234 252 LINKER 12.
 CC FT DOMAIN 253 271 COIL 2A.
 CC FT DOMAIN 272 280 LINKER 2.
 CC FT DOMAIN 281 396 COIL 2B.
 CC FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
 CC FT DOMAIN 444 541 TAIL, SUBDOMAIN B (ACIDIC).
 CC FT CARBOHYD 20 20 O-LINKED (GLCNAC).
 CC FT CARBOHYD 26 26 /FTID-CAR_000128.
 CC FT SITE 381 391 O-LINKED (GLCNAC).
 CC FT CONFLICT 197 202 /FTID-CAR_000129.
 CC FT CONFLICT 399 399 EPTIPE (RECOGNIZED BY IF-SPECIFIC
 CC FT CONFLICT 476 476 MONOCLONAL ANTIBODY).
 CC FT CONFLICT 480 483 GADEA -> KARMSS (IN REF. 2).
 CC FT SEQUENCE 541 AA: 61204 MW: 0D17839AF226918A CRG64;
 CC
 CC Query Match 100.0%; Score 32; DB 1; Length 541;
 CC Best Local Similarity 100.0%; Pred. No. 23;
 CC Matches 7; Conservativity 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC Oy 1 EALEKQL 7
 CC Db 327 EALEKQL 333
 CC
 CC RESULT 2
 CC ID NFL_MOUSE STANDARD; PRT; 542 AA.
 CC AC P08551;
 CC DT 01-AUG-1988 (Rel. 08, Created)
 CC DT 01-OCT-1993 (Rel. 27, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Neurofilament triplet L protein (68 kDa neurofilament protein)
 CC GN NEFL OR NFL.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID=10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE-Brain;
 CC RX MEDLINE=87066433; PubMed=3785173;
 CC RA Lewis S.A., Cowan N.J.;

RT "Anomalous placement of introns in a member of the intermediate
 RT filament multigene family: an evolutionary conundrum";
 RT Mol. Cell. Biol. 6:1529-1534(1986).
 RN [2]
 RN SEQUENCE FROM N.A.
 RN TISSUE-Brain;
 RN MEDLINE=87158637; PubMed=3103856;
 RN Jullien J.-P., Meyer D., Flavell D., Hurst J., Grosfeld F.;
 RT "Cloning and developmental expression of the murine neurofilament
 RT gene family";
 RT Brain Res. 387:243-250(1986).
 RL [3]
 RL SEQUENCE OF 241-542 FROM N.A.
 RL TISSUE-Brain;
 RL MEDLINE=8513134; PubMed=3919033;
 RX Lewis S.A., Cowan N.J.;
 RA "Genetics, evolution, and expression of the 68,000-mol-wt
 RT neurofilament protein: isolation of a cloned cDNA probe";
 RL J. Cell Biol. 100:843-850(1985).
 RN [4]
 RN SEQUENCE OF 1-27 FROM N.A.
 RN MEDLINE=91060592; PubMed=2246261;
 RA Nakahira K., Ikenaka K., Wada K., Tamura T.A., Furutachi T.,
 RA Mikoshiba K.;
 RT "Structure of the 68-kDa neurofilament gene and regulation of its
 RT expression";
 RL J. Biol. Chem. 265:19786-19791(1990).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
 CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
 CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
 CC A CHARGED SCAPFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
 CC OTHER NEURONAL COMPONENTS OR IONS.
 CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X02165; CAB51616.1; -
 CC DR EMBL: M20480; AAA39814.1; -
 CC DR EMBL: M13016; AAA39810.1; -
 CC DR EMBL: M55423; AAA39812.1; -
 CC PIR: A25227; QFMSL.
 CC MGD; MGI:97313; NFL.
 CC Interpro: IPR001664; IF.
 CC Pfam: PF00038; filament; 1.
 CC Prosite: PS00226; IF; 1.
 CC Intermediate filament; Coiled coil; Neurone; Glycoprotein.
 CC INT_MET 0 0
 CC FT DOMAIN 1 92 HEAD.
 CC FT DOMAIN 93 396 ROD.
 CC FT DOMAIN 397 542 TAIL.
 CC FT DOMAIN 93 124 COIL 1A.
 CC FT DOMAIN 125 137 LINKER 1.
 CC FT DOMAIN 138 233 COIL 1B.
 CC FT DOMAIN 234 252 LINKER 12.
 CC FT DOMAIN 253 271 COIL 2A.
 CC FT DOMAIN 272 280 LINKER 2.
 CC FT DOMAIN 281 396 COIL 2B.
 CC FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
 CC FT DOMAIN 444 542 TAIL, SUBDOMAIN B (ACIDIC).
 CC FT CARBOHYD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
 CC FT CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
 CC

```

FT SITE 381 391 EPIPOPE (RECOGNIZED BY IF-SPECIFIC
FT CONFLICT 5 5 MONOCLONAL ANTIBODY).
FT CONFLICT 8 8 Y -> S (IN REF. 1).
FT CONFLICT 64 64 Y -> I (IN REF. 1).
FT CONFLICT 72 72 M -> K (IN REF. 1).
FT CONFLICT 98 98 V -> L (IN REF. 2).
FT CONFLICT 194 194 D -> H (IN REF. 1).
FT CONFLICT 202 202 R -> A (IN REF. 1).
FT CONFLICT 239 239 MISSING (IN REF. 1).
SO SEQUENCE 542 AA; 61448 MW; 8EE9B8C6F0831DBC CRC64;
Query Match 100.0%; Score 32; DB 1; Length 542;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 327 EALEKOL 333

RESULT 3
NFL_HUMAN STANDARD; PRT; 543 AA.
ID P07196; Q16154;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament triplet polypeptide) (NF-L).
GN NEFL OR NEFL OR NF68.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8721423; PubMed=3034332;
RA Julien J.-P., Grosfeld F., Yazdanbakhsh K., Flavell D., Meijer D.,
RA Mushynski W.;
RT "The structure of a human neurofilament gene (NF-L): a unique exon-
RT intron organization in the intermediate filament gene family.";
RL Blochm. Biophys. Acta 909:10-20(1987).
RN [2]
RP SEQUENCE OF 1-10 FROM N.A.
RX MEDLINE=94235564; PubMed=8180132;
RA Pospelov V.A., Pospelova T.V., Julien J.-P.;
RT "Ap-1 and Krox-24 transcription factors activate the neurofilament
RT light gene promoter in P19 embryonal carcinoma cells.";
RL Cell Growth Differ. 5:187-196(1994).
RN [3]
RP VARIANT CMT2E PRO-332.
RX MEDLINE=20307176; PubMed=10841809;
RA Merzlyanova I.V., Perepelov A.V., Polyakov A.V., Stuklov V.F.,
RA Dadaei E.L., Oparin R.B., Petlin A.N., Evgrafov O.V.;
RT "A new variant of Charcot-Marie-Tooth disease type 2 is probably the
RT result of a mutation in the neurofilament-light gene.";
RL Am. J. Hum. Genet. 67:37-46(2000).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- DISEASE: DEFECTS IN NEFL ARE A CAUSE OF CHARCOT-MARIE-TOOTH
CC DISEASE TYPE 2E (CMT2E).
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -1- DATABASE: NAME-INHERITED peripheral neuropathies mutation db;

```

```

CC WWW="http://molgen-www.uia.ac.be/CMTmutations/".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05608; CAA29097.1; -
DR EMBL; S70309; AAD14057.1; -
DR PIR; S07144; S07144.
DR Genew; HGNC:7739; NEFL.
DR MTM; 162280;
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00226; IF; 1.
KM Intermediate filament; Charcot-Marie-Tooth disease.
KW Disease mutation; Charcot-Marie-Tooth disease.
FT INIT_MET 0
FT DOMAIN 1 91 HEAD.
FT DOMAIN 92 396 ROD.
FT DOMAIN 397 543 TAIL.
FT DOMAIN 92 123 COIL_1A.
FT DOMAIN 124 136 LINKER_1.
FT DOMAIN 137 234 COIL_1B.
FT DOMAIN 235 252 LINKER_12.
FT DOMAIN 253 271 COIL_2A.
FT DOMAIN 272 280 LINKER_2.
FT DOMAIN 281 396 COIL_2B.
FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
FT DOMAIN 444 543 TAIL, SUBDOMAIN B (ACIDIC).
FT CARBOHYD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
FT CARBOHYD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
FT SITE 381 391 EPIPOPE (RECOGNIZED BY IF-SPECIFIC
FT SITE 381 391 MONOCLONAL ANTIBODY).
FT FT 0 -> P (IN CMT2E).
FT FT 0 -> P (IN CMT2E).
SO SEQUENCE 543 AA; 61645 MW; 7A0F1AD5BD22F6 CRC64;
Query Match 100.0%; Score 32; DB 1; Length 543;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 327 EALEKOL 333

RESULT 4
NFL_PIG STANDARD; PRT; 548 AA.
ID NFL_PIG
AC P02547;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament triplet polypeptide) (NF-L).
GN NEFL.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=85154583; PubMed=3920075;
RA Geisler N., Plessmann U., Weber K.;
RT "The complete amino acid sequence of the major mammalian
RT neurofilament protein (NF-L).";
RL FEBS Lett. 182:475-478(1985).
RN [2]
RP SEQUENCE OF 1-82 AND 278-548.

```

```

RA RA Gelsele N., Kaufmann E., Flaschert S., Plessmann U., Weber K.;
RT "Neurofilament architecture combines structural principles of
RT intermediate filaments with carboxy-terminal extensions increasing
RT in size between triplet proteins";
RL EMBO J. 2:1295-1302(1983).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC -1- AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
DR DR PIR: A02963; OEPGL.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neurone; Glycoprotein.
FT FT DOMAIN 1 91 HEAD.
FT FT DOMAIN 92 395 ROD.
FT FT DOMAIN 396 548 TAIL.
FT FT DOMAIN 92 123 COIL 1A.
FT FT DOMAIN 124 136 LINKER 1.
FT FT DOMAIN 137 232 COIL 1B.
FT FT DOMAIN 233 251 LINKER 12.
FT FT DOMAIN 252 270 COIL 2A.
FT FT DOMAIN 271 279 LINKER 2.
FT FT DOMAIN 280 395 COIL 2B.
FT FT DOMAIN 396 442 TAIL, SUBDOMAIN A.
FT FT DOMAIN 443 548 TAIL, SUBDOMAIN B (ACIDIC).
FT FT CARBOHD 20 20 O-LINKED (GLCNAC) (BY SIMILARITY).
FT FT CARBOHD 26 26 O-LINKED (GLCNAC) (BY SIMILARITY).
FT FT SITE 380 390 EPITOPE (RECOGNIZED BY IF-SPECIFIC
FT FT MONOCLONAL ANTIBODY).
FT FT UNSURE 322 322 OR K.
FT FT SO SEQUENCE 548 AA; 61940 MW; 83044813637AC739 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 548;
Best Local Similarity 100.0%; Pred. NO. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKQL 7
DB 326 EALEKQL 332
|||||

RESULT 5
NFL_BOVIN STANDARD; PRT; 554 AA.
AC P02548; P79127;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1987 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L) (Micro glutamic acid-rich
DE protein).
GN NEFL.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Brain;
RA Hall W.D., Zhang L., Balin B.J., Sprinkle T.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RP [2]
RP SEQUENCE OF 468-554.

```

```

RX MEDLINE-85154567; PubMed-3884373;
RA Isobe T., Okuyama T.;
RT "Brain micro glutamic acid-rich protein is the C-terminal endpiece of
RT the neurofilament 68-kDa protein as determined by the primary
RT sequence.";
RL FEBS Lett. 182:389-392(1985).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPEECE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAPFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
CC EMBL, U83919; ABA1543.1; -.
CC PIR, A02964; QFBO.
CC InterPro: IPR001664; IF.
CC Pfam: PF00038; filament.1.
CC PROSITE, PS00226; IF; 1.
CC KW Intermediate filament; Coiled coil; Neurope.
CC FT INT MET 0 0
CC FT DOMAIN 1 92 BY SIMILARITY.
CC FT DOMAIN 93 396 HEAD (BY SIMILARITY).
CC FT DOMAIN 397 554 ROD (BY SIMILARITY).
CC FT DOMAIN 93 124 TAIL (BY SIMILARITY).
CC FT DOMAIN 125 137 COIL 1A.
CC FT DOMAIN 138 233 LINKER 1.
CC FT DOMAIN 234 252 COIL 1B.
CC FT DOMAIN 253 271 LINKER 12.
CC FT DOMAIN 272 280 COIL 2A.
CC FT DOMAIN 281 396 LINKER 2.
CC FT DOMAIN 397 443 COIL 2B.
CC FT DOMAIN 444 554 TAIL, SUBDOMAIN A.
CC FT DOMAIN 494 500 TAIL, SUBDOMAIN B (ACIDIC).
CC FT CONFLICT 494 500 MISSING (IN REF. 2).
CC FT CONFLICT 509 509 A -> AEA (IN REF. 2).
CC FT SEQUENCE 554 AA; 62514 MW; D772B81CA2C1C1CA CRC64;
CC -----
Query Match 100.0%; Score 32; DB 1; Length 554;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EALEKQL 7
DB 327 EALEKQL 333
-----
RESULT 6
NFL_COTJA STANDARD; PRT; 555 AA.
AC Q02916;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (Neurofilament light polypeptide) (NF-
DE L).
DE NEFL.
GN Coturnix coturnix japonica (Japanese quail).
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.

```

```

OX NCBI_TaxID-93934;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-93224534; PubMed-846835;
RA Onara O., Gahara Y., Miyake T., Teraoka H., Kitamura T.;
RT "Neurofilament deficiency in quail caused by nonsense mutation in
neurofilament-L gene.";
RL J. Cell Biol. 121:387-395(1993).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTRAIN THREE IF PROTEINS: L, M,
AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM
A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
OTHER NEURONAL COMPONENTS OR IONS.
CC -1- DISEASE: NF-L DEFICIENCY CAUSES THE DISORDER OLIVER.
CC -1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D13223; BAA02504.1;
DR EMBL; D13222; BAA02503.1; ALT_TERM.
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament. 1.
DR PROSITE; PS00226; IF. 1.
KM Intermediate filament; Coiled coil; Neurone.
FT INT_MET 0
FT DOMAIN 1 93 BY SIMILARITY.
FT DOMAIN 94 397 HEAD (BY SIMILARITY).
FT DOMAIN 398 555 ROD (BY SIMILARITY).
FT DOMAIN 94 135 TAIL (BY SIMILARITY).
FT DOMAIN 126 138 COIL 1A.
FT DOMAIN 139 234 COIL 1B.
FT DOMAIN 235 253 LINKER 1.
FT DOMAIN 254 272 LINKER 12.
FT DOMAIN 273 281 COIL 2A.
FT DOMAIN 282 397 LINKER 2.
FT DOMAIN 398 444 COIL 2B.
FT DOMAIN 445 555 TAIL, SUBDOMAIN B (ACIDIC).
SQ SEQUENCE 555 AA; 62282 MW; 9B957ABDBEBA7712 CRC64;

Query Match 100.0%; Score 32; DB 1; Length 555;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKQL 7
Db 328 EALEKQL 334

RESULT 7
DMD_CHICK
ID DMD_CHICK STANDARD; PRT; 3660 AA.
AC P11533;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Dyatrophin.
GN DMD.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

```

```

OC Gallus.
OX NCBI_TaxID-9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Muscle;
RX MEDLINE-89098331; PubMed-3062582;
RA Lemire C., Hellig R., Mandel J.L.;
RT "Nucleotide sequence of chicken dystrophin cDNA.";
RL Nucleic Acids Res. 16:11815-11815(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Muscle;
RX MEDLINE-89210800; PubMed-3072195;
RA Lemire C., Hellig R., Mandel J.L.;
RT "The chicken dystrophin cDNA: striking conservation of the C-terminal
coding and 3' untranslated regions between man and chicken.";
RL EMBL J. 7:4157-4162(1988).
CC -1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE
PLASMA MEMBRANE.
CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIBRIN,
ABP-120, ABP-180, OR BETA-FODRIN).
CC -1- SIMILARITY: CONTAINS 2 CALPONTIN-HOMOLOGY (CH) DOMAINS.
CC -1- SIMILARITY: CONTAINS 22 SPECTRIN REPEATS.
CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13369; CA31746.1;
DR PIR; S02041; S02041.
DR HSP; P46939; IBD.
DR InterPro; IPR001589; Actbind.actuin.
DR InterPro; IPR001715; Calponton-1like.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR002349; WW.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR InterPro; IPR000433; Znf_Z2.
DR Pfam; PF00397; CH. 2.
DR Pfam; PF00435; spectrin. 22.
DR Pfam; PF00569; Z2. 1.
DR PRINTS; PR00403; WMDOMAIN.
DR SMART; SM00033; CH. 2.
DR SMART; SM00150; SPEC. 21.
DR SMART; SM00456; WW. 1.
DR SMART; SM00291; Znf_Z2. 1.
DR PROSITE; PS00019; ACTININ_1.
DR PROSITE; PS00020; ACTININ_2.
DR PROSITE; PS01159; WW_DOMAIN_1.
DR PROSITE; PS01021; CH. 2.
DR PROSITE; PS50020; ACTININ_1.
DR PROSITE; PS50020; WW_DOMAIN_2.
DR PROSITE; PS01357; Znf_Z2. 1.
DR PROSITE; PS01357; Znf_Z2. 1.
DR PROSITE; PS01357; Znf_Z2. 1.
DR PROSITE; PS01357; Znf_Z2. 1.
KM Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;
Repeat; Zinc-finger.
FT DOMAIN 1 244 ACTIN-BINDING.
FT DOMAIN 19 123 CH 1.
FT DOMAIN 138 241 CH 2.
FT REPEAT 341 449 SPECTRIN 1.
FT REPEAT 450 558 SPECTRIN 2.
FT REPEAT 561 669 SPECTRIN 3.
FT REPEAT 721 830 SPECTRIN 4.
FT REPEAT 832 936 SPECTRIN 5.
FT REPEAT 945 1047 SPECTRIN 6.
FT REPEAT 1050 1156 SPECTRIN 7.
FT REPEAT 1159 1265 SPECTRIN 8.
FT REPEAT 1268 1369 SPECTRIN 9.

```



```

FT REPEAT 1470 1570 SPECTRIN 10.
FT REPEAT 1573 1678 SPECTRIN 11.
FT REPEAT 1681 1782 SPECTRIN 12.
FT REPEAT 1879 1981 SPECTRIN 13.
FT REPEAT 2013 2103 SPECTRIN 14.
FT REPEAT 2216 2211 SPECTRIN 15.
FT REPEAT 2472 2574 SPECTRIN 16.
FT REPEAT 2577 2683 SPECTRIN 17.
FT REPEAT 2686 2799 SPECTRIN 18.
FT REPEAT 2802 2904 SPECTRIN 19.
FT REPEAT 2906 2928 SPECTRIN 20.
FT REPEAT 2931 3037 SPECTRIN 21.
FT DOMAIN 3052 3085 SPECTRIN 22.
FT ZN_FING 3304 3351 ZN-TYPE.
FT VARIANT 1171 1171 MISSING.
FT VARIANT 1869 1869 O -> H.
FT VARIANT 1885 1885 K -> R.
SQ SEQUENCE 3660 AA; 422874 MW; 85493DAF6D5B6DA4 CRC64;

```

```

Query Match 100.0%; Score 32; DB 1; Length 3660;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 EALEKOL 7
DB 1084 EALEKOL 1090

```

```

RESULT 8
BAIE_EUBSP STANDARD; PRT; 166 AA.
ID BAIE_EUBSP
AC P19412;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bile acid-inducible operon protein E.
GN BAIE.
OS Eubacterium sp. (strain VPI 12708).
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
CC Clostridium.
CX NCBI_TaxID=29347;
RN (1)
RP MEDLINE=91072253; PubMed=2254270.
RA "Cloning and sequencing of a bile acid-inducible operon from
RT Eubacterium sp. strain VPI 12708."
RL J. Bacteriol. 172:7011-7019(1990).
CC -1- PATHWAY: Bile acid catabolism.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U57489; AAC45413.1; -.
DR PIR: D37844; D37844.
KW Bile acid catabolism.
FT CONFLICT 2 T -> F (IN REF. 1; AA SEQUENCE).
SQ SEQUENCE 166 AA; 19533 MW; 1CBCE86C85ADC35 CRC64;

```

```

Query Match 90.6%; Score 29; DB 1; Length 166;
Best Local Similarity 85.7%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 EALEKOL 7
DB 8 EALEKOL 14

```

```

RESULT 9
RFL_THETH STANDARD; PRT; 354 AA.
ID RFL_THETH
AC P96077;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peptide chain release factor 1 (RF-1);
GN PRFA.
OS Thermus thermophilus.
OC Bacteria; Thermus/Delnococcus group; Deinococci; Thermales;
OC Thermaceae; Thermus.
OX NCBI_TaxID=274;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN-HB8 / ATCC 27634;
RA Ito K., Ebihara K., Nakamura Y.;
RT "Cloning of RF-1 gene from Thermus thermophilus HB8."
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PEPTIDE CHAIN RELEASE FACTOR 1 DIRECTS THE TERMINATION
CC OF TRANSLATION IN RESPONSE TO THE PEPTIDE CHAIN TERMINATION
CC CODONS UAG AND UAA (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC AND MITOCHONDRIAL RELEASE
CC FACTORS FAMILY.
CC -----

```

```

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----

```

```

DR EMBL: D87366; BA13349.1; -.
DR InterPro: IPR005139; PCRF.
DR InterPro: IPR000352; Pep_rel_factor_1.
DR InterPro: IPR004373; Ppfa.
DR Pfam: PF004472; RF-1; 1.
DR Pfam: PF03462; PCRF; 1.
DR TIGRFAMs: TIGR00019; pifa; 1.
DR PROSITE: PS00745; RF_PROK_1; 1.
KW Protein biosynthesis.
SQ SEQUENCE 354 AA; 40092 MW; 4986BD246B3339E CRC64;

```

```

Query Match 90.6%; Score 29; DB 1; Length 354;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 EALEKOL 7
DB 86 EALEKOL 92

```

```

RESULT 10
GABP_MOUSE STANDARD; PRT; 382 AA.
ID GABP_MOUSE
AC O00420;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GA binding protein beta-1 chain (GABP-beta-1 subunit) (GABPB1).
GN GABPB1 OR GABPB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=91433912; PubMed=1876836;
RA Lamarco K., Thompson C.C., Byers B.P., Walton E.M., McKnight S.L.;
RT "Identification of Ets- and notch-related subunits in GA binding

```

```

RT protein".
RL Science 253:789-792(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 5-157.
RX MEDLINE-98128030; PubMed-9461436;
RA Batchelor A.H., Piper D.E., de la Brousse F.C., McKnight S.L.,
RA Wolberger C.;
RT "The structure of GABPalpha/beta: an ETS domain-ankyrin repeat
RT heterodimer bound to DNA.";
RL Science 279:1037-1041(1998).
CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE
CC RICH REPEATS (GA REPEATS).
CC -1- SUBUNIT: HETEROETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.
CC (AC 000421); are produced by alternative splicing. They differ
CC only in the C-terminal extremity.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M74516; AAA53031.1; -
DR PIR: B40858; B40858.
DR PDB: 1ANC; 18-MAR-98.
DR TRANSFAC: T00296; -
DR TRANSFAC: T01403; -
DR MGD: MGI:95611; Gabpbl1.
DR InterPro: IPR002110; ANK.
DR Pfam: PF000023; ank; 4.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 3.
DR PROSITE: PS50088; ANK_REPEAT; 3.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR Transcription regulation; Nuclear protein; Alternative splicing;
DR ANK repeat; Repeat; 3D-structure.
FT REPEAT 5 34 ANK 1.
FT REPEAT 37 66 ANK 2.
FT REPEAT 70 99 ANK 3.
FT REPEAT 103 132 ANK 4.
FT REPEAT 136 166 ANK 5.
SQ SEQUENCE 382 AA; 41258 MW; 0A5FE3F71D42727AF CRC64;

Query Match 90.6%; Score 29; DB 1; Length 382;
Best Local Similarity 85.7%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 335 EALOKOL 341

RESULT 11
GABP_HUMAN
ID GABP_HUMAN STANDARD; PRT; 383 AA.
AC 006547;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GA binding protein beta-1 chain (GABP-beta-1 subunit) (Transcription
DE factor E4TF1-53) (GABP1) (Nuclear respiratory factor-2 subunit beta
DE 2).
GN GABP1 OR GABPB OR E4TF1B.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE-93180783; PubMed-8441384;
RA Watanabe H., Sawada J.-I., Yano K.-I., Yamaguchi K., Goto M.,
RA Handa H.;
RT "cDNA cloning of transcription factor E4TF1 subunits with Ets and
RT notch motifs.";
RL Mol. Cell. Biol. 13:1385-1391(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-95097980; PubMed-7799916;
RA Gueney S., Virdasius J.V., Scarpulla R.C.;
RT "Four structurally distinct, non-DNA-binding subunits of human
RT nuclear respiratory factor 2 share a conserved transcriptional
RT activation domain.";
RL Mol. Cell. Biol. 15:102-111(1995).
CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE
CC RICH REPEATS (GA REPEATS). NECESSARY FOR THE EXPRESSION OF THE
CC ADENOVIRUS E4 GENE.
CC -1- SUBUNIT: HETEROETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.
CC (AC 006545); are produced by alternative splicing.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D13317; BAA02574.1; -
DR HSSP: U13046; AAA65708.1; -
DR TRANSFAC: T01391; -
DR TRANSFAC: T01391; GABP1.
DR Genew; HGNC:4073; GABP1.
DR MIM: 600610; -
DR InterPro: IPR002110; ANK.
DR Pfam: PF000023; ank; 4.
DR PRINTS: PR01415; ANKYRIN.
DR SMART: SM00248; ANK; 3.
DR PROSITE: PS50088; ANK_REPEAT; 3.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR Transcription regulation; Nuclear protein; Alternative splicing;
DR ANK repeat; Repeat.
FT REPEAT 5 34 ANK 1.
FT REPEAT 37 66 ANK 2.
FT REPEAT 70 99 ANK 3.
FT REPEAT 103 132 ANK 4.
FT REPEAT 136 166 ANK 5.
SQ SEQUENCE 383 AA; 41321 MW; 07E7081A60016288 CRC64;

Query Match 90.6%; Score 29; DB 1; Length 383;
Best Local Similarity 85.7%; Pred. No. 70;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 336 EALOKOL 342

RESULT 12
OAT_PLAFD
ID OAT_PLAFD STANDARD; PRT; 414 AA.
AC 007805;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ornithine aminotransferase (EC 2.6.1.13) (Ornithine--oxo-acid
DE amino transferase).
GN OAT.
OS Plasmodium falciparum (isolate CDC / Honduras).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5836;

```

```

RN  (1)
RP  SEQUENCE FROM N.A.
RX  MEDLINE-94088678; PubMed-8264733;
RA  Schmid S.R., Linder P., Reese R.T., Stanley H.A.;
RT  "Characterization of a putative ornithine aminotransferase gene of
RL  Plasmodium falciparum."
CC  Mol. Biochem. Parasitol. 61:311-314(1993).
CC  -1- CATALYTIC ACTIVITY: L-ornithine + a 2-oxo acid - L-glutamate 5-
CC  semialdehyde + an L-amino acid.
CC  -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC  -1- SUBUNIT: HOMOTETRAMER.
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable)
CC  -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC  AMINOTRANSFERASES.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: L15426; AAA16481.1; -.
DR  HSSP: P04181; ZOAT.
DR  InterPro: IPR000954; Aminotran_3.
DR  Pfam: PF00202; aminotran_3; 1.
DR  PROSITE: PS00600; AA_TRANSFERRER_Pyridoxal_phosphate.
KW  Transferase; Aminotransferase; Pyridoxal phosphate.
FT  BINDING 262 262 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ  SEQUENCE 414 AA; 46055 MW; EA2264D1FA26DBF CRC64;

Query Match
Best Local Similarity 90.6%; Score 29; DB 1; Length 414;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  1 EALEKOL 7
DB  163 EALEKEL 189

RESULT 13
TRPE_CLOTH STANDARD: PRT; 494 AA.
ID TRPE_CLOTH STANDARD: PRT; 494 AA.
AC P14953;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Anthranilate synthase component I (EC 4.1.3.27).
GN TRPE.
OS Clostridium thermocellum.
CC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
CC Clostridium.
CC NCBI_TaxID=1515;
CC [1]
CC SEQUENCE FROM N.A.
RX MEDLINE-99278056; PubMed-2732211;
RA Sato S., Nakada Y., Hon-Nami K., Yasui K., Shiratsuchi A.;
RT "Molecular cloning and the nucleotide sequence of the Clostridium
RL thermocellum trpe gene."
CC J. Biochem. 105:362-366(1989).
CC -1- CATALYTIC ACTIVITY: Chorismate + L-glutamine - anthranilate +
CC pyruvate + L-glutamate.
CC -1- PATHWAY: Tryptophan biosynthesis; first step.
CC -1- SUBUNIT: Tetramer of two components I and two components II (BY
CC SIMILARITY).
CC -1- MISCELLANEOUS: COMPONENT I CATALYZES THE FORMATION OF ANTHRANILATE
CC USING AMMONIA RATHER THAN GLUTAMINE. WHEREAS COMPONENT II PROVIDES
CC GLUTAMINE AMINOTRANSFERASE ACTIVITY.
CC -1- SIMILARITY: BELONGS TO THE ANTHRANILATE SYNTHASE COMPONENT I
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL: D00399; BAA00300.1; -.
DR  PIR: JX0065; JX0065.
DR  HSSP: Q06128; IODT.
DR  InterPro: IPR005256; Anth_synthI.
DR  InterPro: IPR000350; Chorismate_bind.
DR  Pfam: PF00425; chorismate_bind; 1.
DR  PRINTS: PR00095; ANTSNTHASEI.
DR  PRODOM: PD000779; Chorismate_bind; 1.
DR  TIGRPFAM: TIGR00564; trpe_most; 1.
KW Tryptophan biosynthesis; Lyase.
SQ SEQUENCE 494 AA; 56020 MW; 32DF1EF22344447D CRC64;

Query Match
Best Local Similarity 90.6%; Score 29; DB 1; Length 494;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  1 EALEKOL 7
DB  323 EALEKEL 329.

RESULT 14
ID Y066_NPVAC STANDARD: PRT; 808 AA.
AC Y066_NPVAC
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 94.0 KDa protein in POL-LEF3 intergenic region.
OS Autographa californica nuclear polyhedrosis virus (ACMNPV).
CC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
CC Nucleopolyhedrovirus.
CC NCBI_TaxID=46015;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN=C6;
CC MEDLINE-94303173; PubMed-8030224;
CC Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
CC "The complete DNA sequence of Autographa californica nuclear
CC polyhedrosis virus."
CC Virology 202:586-605(1994).
CC -1- SIMILARITY: TO CORRESPONDING ORF IN OPANPV AND LDMNPV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL: L22858; AA66696.1; -.
DR  Hypothetical protein.
KW DOMAIN 106 117 POLY-PRO.
FT SEQUENCE 808 AA; 93973 MW; 76A871D2B6633F8A CRC64;
SQ SEQUENCE 808 AA; 93973 MW; 76A871D2B6633F8A CRC64;

Query Match
Best Local Similarity 90.6%; Score 29; DB 1; Length 808;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  1 EALEKOL 7
DB  447 EALEKOL 453

RESULT 15

```

```

AKA9_RABIT
ID AKA9_RABIT STANDARD; PRT; 1087 AA.
AC Q28628;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE A-kinase anchor protein 9 (Protein kinase A anchoring protein 9)
DE (PRKA9) (A-kinase anchor protein 120 kDa) (AKAP 120) (Fragment).
GN AKAP9 OR AKAP120.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Gastric parietal cell;
RX MEDLINE=97220389; PubMed=9148752;
RA Dransfield D.T., Yeh J.L., Bradford A.J., Goldenring J.R.;
RT "Identification and characterization of a novel A-kinase-anchoring
RT protein (AKAP120) from rabbit gastric parietal cells.";
RL Biochem J. 322:801-808(1997)
CC -!- FUNCTION: BINDS TO TYPE II REGULATORY SUBUNITS OF PROTEIN KINASE
CC A. MAY BE A SCAFFOLDING PROTEIN.
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN GASTRIC Parietal CELLS.
CC -!- DOMAIN: RII BINDING SITE, PREDICTED TO FORM AN AMPHIPATHIC HELIX,
CC COULD PARTICIPATE IN PROTEIN-PROTEIN INTERACTIONS WITH A
CC COMPLEMENTARY SURFACE ON THE R-SUBUNIT DIMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U26360; AAC35413.1; ALT_INIT.
KM Coiled coll.
FT NON_TER 1 1
FT DOMAIN 559 572 PKA-RII SUBUNIT BINDING DOMAIN.
FT DOMAIN 3 461 COILED COIL (POTENTIAL).
FT DOMAIN 614 773 COILED COIL (POTENTIAL).
FT NON_TER 1087 1087
SQ SEQUENCE 1087 AA; 124756 MW; 9D916BEOCA89FF02 CRC64;

Query Match 90.6%; Score 29; DB 1; Length 1087;
Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 EALEKOL 7
Db 54 QALEKOL 60

```

Search completed: November 13, 2002, 13:16:51
 Job time : 6.80851 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 22.4894 Seconds
(without alignments)
64.134 Million cell updates/sec

Title: US-09-856-086-3

Perfect score: 32

Sequence: 1 EALRKQL 7

Scoring table: BLOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL.21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rv1rus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	32	100.0	309	17	Q8TSK1	Q8TSK1 methanosaarc
2	32	100.0	386	4	Q8RCR7	Q8RCR7 homo sapien
3	32	100.0	727	12	Q91F05	Q91F05 cydia pomon
4	30	93.8	79	16	Q8RB86	Q8RB86 thermotomaer
5	30	93.8	203	2	Q9A015	Q9A015 streptococc
6	30	93.8	243	2	Q87187	Q87187 streptococc
7	30	93.8	384	2	Q9S0S6	Q9S0S6 streptococc
8	30	93.8	384	2	Q9RPS9	Q9RPS9 streptococc
9	30	93.8	384	2	Q9ALW5	Q9ALW5 streptococc
10	30	93.8	384	2	Q9AFH0	Q9AFH0 streptococc
11	30	93.8	384	2	Q937T1	Q937T1 streptococc
12	30	93.8	451	16	Q9KER7	Q9KER7 bacillus ha
13	30	93.8	459	3	Q94729	Q94729 schizosacch
14	29	90.6	135	2	Q8VHM2	Q8VHM2 pseudomonas
15	29	90.6	238	10	Q9SMO5	Q9SMO5 oryza sativ
16	29	90.6	246	10	Q9LEH9	Q9LEH9 hordeum vul

17	29	90.6	249	10	Q04063	Q04063 oryza sativ
18	29	90.6	249	10	Q24234	Q24234 oryza sativ
19	29	90.6	253	16	Q51655	Q51655 borrelia bu
20	29	90.6	259	10	Q24229	Q24229 oryza sativ
21	29	90.6	280	10	Q9LHD4	Q9LHD4 arabidopsis
22	29	90.6	284	4	Q9UFL2	Q9UFL2 homo sapien
23	29	90.6	337	17	Q59045	Q59045 pyrococcus
24	29	90.6	378	5	Q9VM45	Q9VM45 pyrococcus
25	29	90.6	382	11	Q9QVY2	Q9QVY2 mus sp. gab
26	29	90.6	383	11	Q91Y20	Q91Y20 mus musculu
27	29	90.6	395	4	Q12940	Q12940 homo sapien
28	29	90.6	395	5	Q960D9	Q960D9 homo sapien
29	29	90.6	400	17	Q8UZN8	Q8UZN8 pyrococcus
30	29	90.6	414	5	Q9G0I3	Q9G0I3 plasmodium
31	29	90.6	489	11	Q93KB2	Q93KB2 mus musculu
32	29	90.6	502	4	Q9UDV2	Q9UDV2 homo sapien
33	29	90.6	551	3	Q43090	Q43090 schizosacch
34	29	90.6	654	2	Q59656	Q59656 plectonema
35	29	90.6	721	5	Q95R25	Q95R25 drosophila
36	29	90.6	727	4	Q75127	Q75127 homo sapien
37	29	90.6	776	5	Q9VY40	Q9VY40 drosophila
38	29	90.6	805	12	Q92431	Q92431 bombyx mori
39	29	90.6	1074	17	Q8TRL1	Q8TRL1 methanosaarc
40	29	90.6	1078	2	Q49529	Q49529 mycoplasma
41	29	90.6	1079	2	Q49555	Q49555 mycoplasma
42	29	90.6	1249	5	Q95RV1	Q95RV1 drosophila
43	29	90.6	1262	5	Q9N3I1	Q9N3I1 leishmania
44	29	90.6	1463	5	Q96GT0	Q96GT0 strongyloce
45	29	90.6	1504	5	Q9VY86	Q9VY86 drosophila

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	309 AA.
1	Q8TSK1			
AC	Q8TSK1	01-JUN-2002 (TRENBLREL. 21, Created)		
DT	01-JUN-2002 (TRENBLREL. 21, Last sequence update)			
DT	01-JUN-2002 (TRENBLREL. 21, Last annotation update)			
DE	Hypothetical protein MA0795.			
GN	MA0795.			
OS	Methanosaarcina acetivorans.			
OC	Archaea; Euryarchaeota; Methanococci; Methanosaarcinales;			
OC	Methanosaarcinaceae; Methanosaarcina.			
OX	NCBI_TaxID=2214;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C2A / ATCC 35395 / DSM 2834.			
RX	MEDLINE=21929760; PubMed=11932238;			
RA	Galagan J.E., Nusbaum C., Roy A., Endlitz M.G., MacDonald P.,			
RA	FitzHugh W., Calvo S., Engels R., Smitrov S., Alnoor D., Brown A.,			
RA	Allen N., Naylor J., Stange-Thomann N., Deatellano K., Johnson R.,			
RA	Linton L., McEwan P., McKernan K., Talamas J., Turrell A., Ye W.,			
RA	Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,			
RA	Hedderich R., Ingram-Smith C., Kuestner H.C., Krzycki J.A.,			
RA	Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,			
RA	Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,			
RA	Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,			
RA	Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,			
RA	Metcalfe W.W., Birren B.			
RT	"The genome of Methanosaarcina acetivorans reveals extensive metabolic			
RL	and physiological diversity."			
DR	Genome Res. 12:532-542(2002).			
EMBL	AE010741; AA04234.1; .			
KW	Hypothetical protein; Complete proteome.			
SO	SEQUENCE 309 AA; 35362 MW; D25EABFC07FEF6 CMC64;			
Query Match	Score 32; DB 17; Length 309;			
Best Local Similarity	100.0%; Pred. No. 68;			
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			

OY 1 EALEKOL 7
 Db 78 EALEKOL 84

RESULT 2

08TCR7 PRELIMINARY; PRT; 386 AA.
 AC 08TCR7;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical 43.8 kDa protein (Fragment).
 GN DKEF61K0922.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=AMYGDALA;
 RA Wambutt R., Heubner D., Mewes H.W., Well B., Wiemann S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL713644; CAD28456.1; -
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 386 AA; 43779 MW; F293388B200C7B65 CRC64;

Query Match 100.0%; Score 32; DB 4; Length 386;
 Best Local Similarity 100.0%; Pred. No. 83;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
 Db 170 EALEKOL 176

RESULT 3

091F05 PRELIMINARY; PRT; 727 AA.
 AC 091F05;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE ORF50. similar to XCGV ORF47.
 GN ORF50.
 OS Cydia pomonella granulosis virus (CPGV) (Cydia pomonella
 OS granulovirus)
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
 OX NCBI_TaxID=8289;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=9318168; PubMed=8445726;
 RA Crook N.E., Clem R.J., Miller L.K.;
 RT "An Apoptosis-Inhibiting baculovirus gene with a zinc finger-like
 motif."
 RT J. Virol. 67:2168-2174(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=96207404; PubMed=8615018;
 RA Theilmann D.A., Chantler J.R., Stewart S., Filippen H.T., Viak J.M.,
 RA Crook N.E.;
 RT "Characterization of a highly conserved baculovirus structural protein
 that is specific for occlusion-derived virions."
 RT Virology 218:148-158(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MEXICAN 1;
 RX MEDLINE=97380577; PubMed=9237352;
 RA Kang W., Crook N.E., Winstanley D., O'Reilly D.R.;
 RT "Complete sequence and transposon mutagenesis of the BamHI J fragment

RT of Cydia pomonella granulosis virus.",
 RL Virus genes 14:131-136(1997).

RESULT 4

08RB86 PRELIMINARY; PRT; 79 AA.
 AC 08RB86;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical protein TTE0936.
 GN TTE0936.
 OS Thermomicrobacter tengcongensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Thermomicrobacteriales; Thermomicrobacteriaceae; Thermomicrobacter.
 OX NCBI_TaxID=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4T / JCM11007;
 RX MEDLINE=21992816; PubMed=11997336;
 RA Bao O., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of T. tengcongensis genome."
 RT Genome Res. 12:689-700(2002).
 DR EMBL: AE013059; AM24192.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 79 AA; 9078 MW; B3FC525B10E95FF9 CRC64;

Query Match 100.0%; Score 32; DB 12; Length 727;
 Best Local Similarity 100.0%; Pred. No. 1,5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
 Db 211 EALEKOL 217

RESULT 5

09A015 PRELIMINARY; PRT; 203 AA.
 AC 09A015;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE NEUC. (Fragment).
 GN NEUC.
 OS Streptococcus agalactiae.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID-1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miyake K., Watanabe M., Iijima S.;
 RT "cpd of Streptococcus agalactiae type 1b shows beta-1,3-
 galactosyltransferase activity."
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB050723; BAB21606.1; -
 FT NON_TER 203 203
 SQ SEQUENCE 203 AA; 22989 MW; EDPBA75CD02B7CDA CRC64;
 Query Match
 Best Local Similarity 93.8%; Score 30; DB 2; Length 203;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EALEKOL 7
 DB 190 EALEKOL 196

RESULT 6
 O87187
 ID 087187 PRELIMINARY; PRT; 243 AA.
 AC 087187;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DE 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
 DE Neuc protein (fragment).
 GN NEUC.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID-1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamamoto S., Miyake K., Iijima S.;
 RT Identification and Characterization of cps (capsular polysaccharide)
 RT genes from Streptococcus agalactiae Type Ia."
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB017355; BAA33734.1; -
 FT NON_TER 243 243
 SQ SEQUENCE 243 AA; 27429 MW; 4D2D31E775A4E08F CRC64;
 Query Match
 Best Local Similarity 93.8%; Score 30; DB 2; Length 243;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EALEKOL 7
 DB 190 EALEKOL 196

RESULT 7
 O9S0S6
 ID O9S0S6 PRELIMINARY; PRT; 384 AA.
 AC O9S0S6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Neuc.
 GN NEUC.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID-1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98395021; PubMed-10464185;
 RA Yamamoto S., Miyake K., Koike Y., Watanabe M., Machida Y., Ohta M.,
 RA Iijima S.;
 RT "Molecular characterization of type-specific capsular polysaccharide

RT biosynthesis genes of Streptococcus agalactiae type Ia.";
 RL J. Bacteriol. 181:5176-5184(1999).
 DR EMBL: AB028896; BAA83288.1; -
 DR InterPro: IPR001064; Crystalin.
 DR InterPro: IPR000169; Shprot_acsite.
 DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
 DR PROSITE: PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
 SQ SEQUENCE 384 AA; 42922 MW; 3293FA83F135BF8B CRC64;
 Query Match
 Best Local Similarity 93.8%; Score 30; DB 2; Length 384;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EALEKOL 7
 DB 190 EALEKOL 196

RESULT 8
 O9RPB9
 ID O9RPB9 PRELIMINARY; PRT; 384 AA.
 AC O9RPB9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Neuc.
 GN NEUC.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID-1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-COH1;
 RX MEDLINE-20372630; PubMed-10913080;
 RA Chaffin D.O., Beres S.B., Yim H.H., Rubens C.E.;
 RT "The Serotype of Type Ia and Iii Group B Streptococci Is Determined by
 RT the Polymerase Gene within the Polyclonistic Capsule Operon."
 RL J. Bacteriol. 182:4466-4477(2000).
 DR EMBL: AF168833; AAD53075.1; -
 DR InterPro: IPR000169; Shprot_acsite.
 DR PROSITE: PS00639; THIOI_PROTEASE_HIS; UNKNOWN_1.
 SQ SEQUENCE 384 AA; 43018 MW; A190AA39E0F2AB94 CRC64;
 Query Match
 Best Local Similarity 93.8%; Score 30; DB 2; Length 384;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 EALEKOL 7
 DB 190 EALEKOL 196

RESULT 9
 O9ALM5
 ID O9ALM5 PRELIMINARY; PRT; 384 AA.
 AC O9ALM5;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Putative UDP-N-acetylglucosamine-2-epimerase Neuc.
 GN NEUC.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID-1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NT6;
 RA McKinnon K., Chaffin D.O., Rubens C.E.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF337958; AAK11671.1; -
 DR InterPro: IPR000169; Shprot_acsite.

DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN.1.
 SQ SEQUENCE 384 AA; 42952 MW; D39A21E40F8D4D CRC64;
 Query Match 93.8%; Score 30; DB 2; Length 384;
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
 ||:||||
 Db 190 EALEKOL 196

RESULT 10

ID 09AFH0 PRELIMINARY; PRT; 384 AA.
 AC 09AFH0;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Neuc.
 GN NEUC.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID=1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CNCYC 1/82;
 RA McKinnon K., Chaffin D.O., Rubens C.E.;
 RT "Streptococcus agalactiae type V polysaccharide synthesis operon
 complete sequence."
 RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF349539; AAK29662.1;
 DR InterPro: IPR000169; SHProt. acslte.
 DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN.1.
 SQ SEQUENCE 384 AA; 42942 MW; 7D995E2F1F81A50 CRC64;

Query Match 93.8%; Score 30; DB 2; Length 384;
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
 ||:||||
 Db 190 EALEKOL 196

RESULT 11

ID 093T11 PRELIMINARY; PRT; 384 AA.
 AC 093T11;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Putative UDP-N-acetylglucosamine-2-epimerase Neuc.
 GN NEUC.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Streptococcus.
 OX NCBI_TaxID=1311;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CNCYC 1/82;
 RA McKinnon K., Chaffin D.O., Rubens C.E.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR InterPro: IPR000169; SHProt. acslte.
 DR PROSITE; PS00639; THIOI_PROTEASE_HIS; UNKNOWN.1.
 SQ SEQUENCE 384 AA; 43004 MW; 21486868A868E24 CRC64;

Query Match 93.8%; Score 30; DB 2; Length 384;
 Best Local Similarity 85.7%; Pred. No. 2.2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
 ||:||||
 Db 190 EALEKOL 196

RESULT 12

ID 09KER7 PRELIMINARY; PRT; 451 AA.
 AC 09KER7;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Denosylmethionine-8-amino-7-oxononanoate aminotransferase.
 GN BH0782.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 halodurans and genomic sequence comparison with Bacillus subtilis."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001509; BAB04501.1;
 DR HSSP; P04181; 20XT.
 DR InterPro: IPR000954; AminoTran_3.
 DR Pfam; PF00202; aminoTran_3; 1.
 DR TIGRFAMs; TIGR00508; bioA; 1.
 DR PROSITE; PS00600; AA_TRANSFERRER; Complete proteome.
 DR TRANSFERRER; AminoTranTransferase; Complete proteome.
 SQ SEQUENCE 451 AA; 50488 MW; E0FD58D46345D476 CRC64;

Query Match 93.8%; Score 30; DB 16; Length 451;
 Best Local Similarity 85.7%; Pred. No. 2.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 EALEKOL 7
 ||:||||
 Db 68 EALEKOL 74

RESULT 13

ID 094729 PRELIMINARY; PRT; 459 AA.
 AC 094729;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical protein Cl604.17c in chromosome II.
 GN SPBC1604.17C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Ralander M.A., Lyne M., Lyne R., Stewart A.,
 Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 Brooks K., Brown D., Brown R., Chillingworth T., Churcher C.M.,
 Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Saiton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkseert G., Aert R., Robben J., Grymonprez B.,
RA Wiltens I., Vanstreels E., Rieger M., Schaefer M., Muller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzom K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
RA Dage R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe.",
RT Nature 415:871-880(2002).
RN [2]
RP REVISIONS TO N-TERMINUS.
RA Beck A., Reinhardt R., Lyne M., Rajandream M.A., Barrell B.G.,
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL034433; CAA22350.2; -
KW Hypothetical protein; Colled coil.
FT DOMAIN 4
SQ SEQUENCE 459 AA; 5388 MW; 6A1B5D30ABF3616E CRC64;

Query Match 93.8%; Score 30; DB 3; Length 459;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 21 EALEKOL 27

RESULT 14
Q8VMH2 PRELIMINARY; PRT; 135 AA.
AC Q8VMH2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 15.4 kDa protein.
OS Pseudomonas putida.
OG Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID-303;
RN [1]
RP SEQUENCE FROM N.A.
RA Greated A., Lambertson L., Williams P.A., Thomas C.M.,
RT "Complete nucleotide sequence of IncP-9 plasmid pMW0.",
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ344068; CAC68842.1; -
DR InterPro: IPR000551; HTH_MERR.
DR Pfam: PF00376; merr; 1.
DR PRINTS: PR00040; HTHMERR.
DR SMART: SM00422; HTH_MERR; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 135 AA; 15402 MW; D5298D5FE312D483 CRC64;

Query Match 90.6%; Score 29; DB 2; Length 135;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 99 EALEKOL 105

RESULT 15
Q9SMQ5 PRELIMINARY; PRT; 238 AA.
ID Q9SMQ5

AC Q9SMQ5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE MADS-box protein FDRMADS1 (Fragment).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID-4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GUANGDAI 4;
RA Jia H., Cong B., Shao J., Sun C.,
RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF141966; MAD38370.1; -
DR HSP; P11831; 1SR5.
DR InterPro: IPR002487; TF_Rbox.
DR InterPro: IPR002100; TF_MADSbox.
DR Pfam: PF01486; K-box; 1.
DR Pfam: PF00319; SRP-TR; 1.
DR PRINTS: PR00404; MADSDOMAIN.
DR SMART: SM00432; MADS; 1.
DR PROSITE: PSS00066; MADS_BOX_2; 1.
FT NON_TER 1
SQ SEQUENCE 238 AA; 27376 MW; 405E797E5B1152C6 CRC64;

Query Match 90.6%; Score 29; DB 10; Length 238;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EALEKOL 7
DB 117 EALEKOL 123

Search completed: November 13, 2002, 13:16:07
Job time : 24.4894 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OK protein - protein search, using sw model

Run on: November 13, 2002, 13:13:25 : Search time 34.7021 Seconds
(without alignments)
26.879 Million cell updates/sec

Title: US-09-856-086-4

Perfect score: 36

Sequence: 1 ELEDKON 7

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: A_Geneseq_101002.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	7	21	Test antigen #4 fo
2	36	100.0	441	19	Human neurofilamen
3	36	100.0	543	22	Human polyprotein
4	32	88.9	357	23	Herbicidally activ
5	32	88.9	587	21	Arabidopsis thalia
6	32	88.9	624	21	Arabidopsis thalia
7	32	88.9	694	21	Arabidopsis thalia
8	31	86.1	103	21	Human secreted pro
9	31	86.1	180	22	Drosophila melanog
10	31	86.1	298	23	Streptococcus poly

11	31	86.1	321	23	ABP28875
12	30	83.3	400	21	AA07761
13	30	83.3	548	22	ABP27724
14	30	83.3	1713	23	ABP58422
15	30	83.3	2701	17	AA09988
16	29	80.6	25	23	ABG62497
17	29	80.6	158	18	AAW23420
18	29	80.6	159	22	AAU33822
19	29	80.6	159	22	AAU36816
20	29	80.6	189	22	AAU37159
21	29	80.6	199	22	AAU37549
22	29	80.6	196	23	ABP39756
23	29	80.6	229	20	AAU20137
24	29	80.6	232	21	AAU25752
25	29	80.6	274	20	AAU20136
26	29	80.6	276	21	AAU25751
27	29	80.6	278	22	AAU31524
28	29	80.6	278	22	AAU80125
29	29	80.6	278	22	AAU80125
30	29	80.6	313	21	AAU25750
31	29	80.6	327	21	AAU25750
32	29	80.6	346	22	ABP71575
33	29	80.6	441	21	AAU46685
34	29	80.6	441	21	AAU46685
35	29	80.6	457	21	AAU46684
36	29	80.6	757	19	AAU60486
37	29	80.6	771	19	AAU60590
38	29	80.6	801	22	AAU94270
39	29	80.6	819	23	ABP27740
40	29	80.6	850	23	ABP93031
41	29	80.6	859	22	ABP71182
42	29	80.6	1351	22	ABG04063
43	29	80.6	3574	22	AAU14347
44	29	80.6	4767	22	ABU11532
45	28	77.8	41	22	ABU03166

ALIGNMENTS

RESULT 1
AAB05928 standard; peptide: 7 AA.
ID AAB05928;
AC AAB05928;
XX 17-OCT-2000 (first entry)
XX Test antigen #4 for spongiform and demyelinating disease diagnosis.
XX Human: cow; myelin; myelin neurofilament; immunogen; antigen;
XX bovine spongiform encephalopathy; BSE; multiple sclerosis;
XX Creutzfeldt-Jacob disease; CJD; demyelinating disease; diagnostic test.
XX Bos taurus.
XX Homo sapiens.
XX WO200031545-A1.
XX 02-JUN-2000.
XX 25-NOV-1999; 99WO-GB03936.
XX 26-NOV-1998; 98GB-0025948.
XX (UNLO) KING'S COLLEGE.
XX Ebringer A;
XX WPI; 2000-400194/34.
XX Diagnosing spongiform or demyelinating disease in vertebrates such as
XX bovine spongiform encephalopathy and Creutzfeldt-Jacob disease comprises

PT assaying a biological sample for myelin and/or myelin neurofilament
 XX antibodies -
 PS Claim 5; Page 2; 16pp; English.
 XX
 CC The present peptide may be used as a test antigen in a kit for diagnosing
 CC Spongiform or demyelinating disease in vertebrates, including bovine
 CC Creutzfeldt-Jacob disease (CJD). Peptides used in the kit are antigenic
 CC components of myelin or myelin neurofilaments. Biological samples
 CC are assayed for antibodies, especially IgA antibodies, which bind to
 CC myelin and/or myelin neurofilaments or an antigenic peptide, such as the
 CC present sequence. Any reading in excess of two standard deviations of
 CC the healthy controls would indicate a positive response.
 SQ Sequence 7 AA:
 Query Match 100.0%; Score 36; DB 21; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ELEDKON 7
 Db 1 ELEDKON 7
 RESULT 2
 ID AAY20612 standard; Protein; 441 AA.
 XX AAY20612;
 XX
 XX 22-JUL-1999 (first entry)
 DE Human neurofilament-L wild type protein fragment 2.
 XX
 KM Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KM frameshift mutation; age-related disease; neurodegenerative disorder;
 KM Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KM Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KM diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KM ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KM neurofilament-F; presenilin 1; presenilin II; cellular tumour antigen;
 KM glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KM bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMBP-C; NSP-A;
 KM high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Homo sapiens.
 XX
 PN W09845322-A2.
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98MO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJCSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 DR WPI; 1998-609901/51.
 DR N-PSDB; AAX75758.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 XX RNA
 XX Disclosure; Figure 7; 258pp; English.

XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMBP-C) and neuroendocrine specific protein A.
 XX
 SQ Sequence 441 AA:
 Query Match 100.0%; Score 36; DB 19; Length 441;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ELEDKON 7
 Db 394 ELEDKON 400
 RESULT 3
 ID AAM93466 standard; Protein; 543 AA.
 XX AAM93466;
 XX
 AC AAM93466;
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polypeptide, SEQ ID NO: 3132.
 XX
 KM Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX
 OS Homo sapiens.
 XX
 PN EP1130094-A2.
 PD 05-SEP-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 99JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELT-) HELIX RES INST.
 XX
 PI Oka T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 DR N-PSDB; AAK94387.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 3132; 1380pp + sequence listing; English.
 CC
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143277.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143824.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145099.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145146.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.9%; Score 32; DB 21; Length 587;
Best Local Similarity 85.7%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ELEDKON 7
Db 166 EFEDKON 172

RESULT 6
AAG42402
ID AAG42402 standard; Protein; 624 AA.
XX
AC AAG42402;

XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 52878.
DE Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX EPI033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 03-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139432.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-01446814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148368.
PR 17-AUG-1999; 99US-0148175.
PR 18-AUG-1999; 99US-0148426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152353.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 88.9%; Score 32; DB 21; Length 624;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 ELEDKON 7
DB 203 EEDKON 209

RESULT 7
AAG42401
ID AAG42401 standard; Protein: 694 AA.
XX AAG42401;
XX 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 52877.
DE Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
KW

KM termination sequence.
XX Arabidopsis thaliana.
OS EPI033405-A2.
XX 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134376.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135323.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.

```

PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

```

```

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 88.9%; Score 32; DB 21; Length 694;
 Best Local Similarity 85.7%; Pred. No. 4.5e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
 DB 273 EFDKON 279

RESULT 8
 AAG02130
 ID AAG02130 standard; Protein; 103 AA.

AC AAG02130;
 DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 6211.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 gene therapy; chromosome mapping.

OS Homo sapiens.

XX EPI033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.
 XX 26-FEB-1999; 99US-0122487.
 XX (GEST) GENSET.
 PA Dumas Mline Edwards J, Duclert A, Giordano J;
 XX WPI; 2000-500381/45.
 DR N-PSDB; AAC02136.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13: SEQ ID 6211; 71pp + CD-ROM; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 103 AA;
 XX
 Query Match 86.1%; Score 31; DB 21; Length 103;
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ELEDKON 7
 | | | | |
 Db 89 EXEDKON 95
 XX
 RESULT 9
 ABB68503
 ID ABB68503 standard; Protein; 180 AA.
 XX
 AC ABB68503;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 32301.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical.
 KM
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL12606.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure: SEQ ID NO 32301; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (AB1857737-AB1872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 180 AA;
 XX
 Query Match 86.1%; Score 31; DB 22; Length 180;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LEDKON 7
 | | | | |
 Db 23 LEDKON 28
 XX
 RESULT 10
 ABP30505
 ID ABP30505 standard; Protein; 298 AA.
 XX
 AC ABP30505;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 10186.
 XX
 XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KM group A streptococcus; Streptococcus pyogenes; antibacterial;
 KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
 OS
 XX Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Maignan V, Margalit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 DR WPI; 2002-352536/38.
 DR N-PSDB; ABN71136.
 XX
 PF New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 4148; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in

CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
CC
SQ Sequence 298 AA;
Query Match 86.1%; Score 31; DB 23; Length 298;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LEKON 7
|||||
DB 274 LEKON 279
RESULT 11
ABP28875
ID ABP28875 standard; Protein: 321 AA.
XX
AC ABP28875;
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 6926.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Maignant V, Margarit Ros YI, Grandl G, Fraser C,
PI Tecteln H;
XX
DR MPI; 2002-35236/38.
DR N-PSDB; ABN69506.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1; Page 3855; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for

CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
CC
SQ Sequence 321 AA;
Query Match 86.1%; Score 31; DB 23; Length 321;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 LEKON 7
|||||
DB 297 LEKON 302
RESULT 12
AAB07761
ID AAB07761 standard; Protein: 400 AA.
XX
AC AAB07761;
DT 07-NOV-2000 (first entry)
XX
DE Amino acid sequence of a BASB047 polypeptide.
XX
KW BASB047; BASB054; BASB068; BASB069; vaccine; bacteremia; meningitis;
KW Neisseria meningitidis disease; upper respiratory tract infection;
KW invasive bacterial disease.
XX
OS Neisseria meningitidis.
XX
PN WO200043519-A2.
XX
PD 27-JUL-2000.
XX
PF 19-JAN-2000; 2000WO-EP00428.
XX
PR 22-JAN-1999; 99GB-0001368.
PR 28-JAN-1999; 99GB-0001944.
PR 29-JAN-1999; 99GB-0002086.
PR 15-FEB-1999; 99GB-0003417.
PR 16-FEB-1999; 99GB-0003535.
XX
PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J;
XX
DR MPI; 2000-505839/45.
DR N-PSDB; AAA59347.
XX
PT Neisseria meningitidis BASB047, BASB054, BASB068, and BASB069 proteins,
PT useful for treating N. meningitidis infections, bacteremia, and
PT meningitis -
XX
PS Claim 3; Page 93-94; 103pp; English.
XX
CC The specification describes Neisseria meningitidis polypeptides
CC designated BASB047, BASB054, BASB068, and BASB069. The polynucleotide
CC sequences can be used to create a vector to transform a host cell.
CC The host cell can be used to produce the polypeptide. The polynucleotides
CC and polypeptides can be used in vaccine compositions. The
CC polynucleotides, polypeptides, and antibodies directed against the
CC polypeptides can be used in compositions for preparation of medicaments.
CC The antibodies can also be used in a composition for treating humans
CC with Neisseria meningitidis disease. The diseases that can be treated

CC Include upper respiratory tract infection, and invasive bacterial
CC diseases such as bacteremia and meningitis. The nucleic acid sequences
CC can be used as probes in the diagnosis of *Neisseria meningitidis* disease.
CC The present sequence represents a BAS047 polypeptide.
XX

SO Sequence 400 AA;

Query Match 83.3%; Score 30; DB 21; Length 400;
Best Local Similarity 71.4%; Pred. No. 6.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
:|||||
Db 250 QLDKON 256

RESULT 13

ABP27724
ID ABP27724 standard; Protein; 548 AA.

AC ABP27724;

DT 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 4624.

XX Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
XX group A streptococcus; Streptococcus pyogenes; antibacterial;
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.

OS Streptococcus pyogenes.

PN WO200234771-A2.

PD 02-MAY-2002.

PF 29-OCT-2001; 2001WO-GB04789.

PR 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

PA (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

PI Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettein H;

XX WPI: 2002-352536/38.

DR N-PSDB; ABN68355.

XX New Streptococcus protein for the treatment or prevention of infection
XX or disease caused by Streptococcus bacteria, such as meningitis, and
XX for detecting a compound that binds to the protein -
XX

PS Claim 1; Page 3624; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (1), nucleic acids encoding (1), ABN6044-ABN71526 and
CC antibodies that bind (1) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC streptococcus bacteria, particularly *S. agalactiae* and *S. pyogenes*.
CC Nucleic acids encoding (1) are used to detect Streptococcus in a
CC biological sample. (1) is used to determine whether a compound binds to
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (1) may be used to recombinantly produce (1) and may be
CC used in gene therapy. Antibodies to (1) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying

CC Streptococcus proteins.

SO Sequence 548 AA;

Query Match 83.3%; Score 30; DB 23; Length 548;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKQ 6
:|||||
Db 134 ELEDKQ 139

RESULT 14

ABB58422
ID ABB58422 standard; Protein; 1713 AA.

AC ABB58422;

DT 26-MAR-2002 (first entry)

XX *Drosophila melanogaster* polypeptide SEQ ID NO 2058.

XX *Drosophila*; developmental biology; cell signalling; insecticide;
XX pharmaceutical.

OS *Drosophila melanogaster*.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PA Venter JC, Adams M, Li PWD, Myers EW;

DR WPI: 2001-656860/75.

DR N-PSDB; ABL02525.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from *Drosophila* and for elucidating cell signalling and cell-cell
XX interactions -
XX

PS Disclosure; SEQ ID NO 2058; 21pp + Sequence listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB5737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 1713 AA;

Query Match 83.3%; Score 30; DB 22; Length 1713;
Best Local Similarity 85.7%; Pred. No. 2.7e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDKON 7
:|||||
Db 635 ELEDKON 641

RESULT 15
AAB00050

ID AAR99850 standard; protein; 2701 AA.
XX

AC AAR99850;
VV

DT 23-OCT-1996 (first entry)
XX

Human type 3 inositol triphosphate receptor.

IP₃; inositol triphosphate; receptor; calcium release inhibitor.

OS Homo sapiens.
XX

PN JP08134097-A.
XX

PD 28-MAY-1996
XX

PF 02-NOV-1994; 94JP-0293933.
XX

PR 02-NOV-1994; 94JP-0293933.
XX

PA (SOSE-) SOSEI KK.
XY

WPI; 1996-306573/31.

PT	Human type 2 inositol tri:phosphate receptor - used to study
PT	activity of inositol tri:phosphate as a calcium release inhibitor
xx	

PS Claim 1; Page 4-12; 12pp; Japanese.
xy

CC The present sequence represents the amino acid sequence of the human
CC type 2 inositol triphosphate (IP3) receptor. The receptor is useful as
CC a reagent for research on the activity and the mechanism of IP3 and as a
CC Calcium release inhibitor.

SQ Sequence 2701 AA;

Query Match	83.3%	Score 30	DB 17	Length 2701
Best Local Similarity	85.7%	Score 34	DB 17	Length 2701

```

Best Local Similarity 85.7%; Pred. No. 4.2e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ELEDKON 7

Db 97 ELEQKQN 103

Search completed: November 13, 2002, 13:22:09
Job time : 36.7021 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2002, 13:13:25 ; Search time 10.1277 Seconds
(Without alignments)
20.336 Million cell updates/sec

Title: US-09-856-086-4

Perfect score: 36

Sequence: 1 ELEDKCN 7

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued_Patents_AA*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	160	2	US-08-726-306A-183
2	29	80.6	158	2	US-08-618-911-6
3	29	80.6	156	4	US-09-134-001C-4601
4	29	80.6	156	4	US-09-134-001C-4601
5	28	77.8	133	3	US-08-742-753-2
6	28	77.8	133	3	US-08-966-318-5
7	28	77.8	133	4	US-09-216-619-5
8	28	77.8	133	4	US-08-887-534A-2
9	28	77.8	133	4	US-08-887-534A-7
10	28	77.8	133	4	US-08-533-306A-2
11	28	77.8	133	4	US-08-887-534A-30
12	28	77.8	133	4	US-08-533-306A-6
13	28	77.8	133	4	US-08-742-923A-6
14	28	77.8	133	4	US-08-533-306A-4
15	28	77.8	133	4	US-08-742-923A-6
16	28	77.8	133	4	US-08-533-306A-4
17	28	77.8	133	4	US-08-742-923A-6
18	28	77.8	133	4	US-08-533-306A-2
19	28	77.8	133	4	US-08-887-534A-30
20	28	77.8	133	4	US-08-533-306A-6
21	28	77.8	133	4	US-08-742-923A-6
22	28	77.8	133	4	US-08-533-306A-4
23	28	77.8	133	4	US-08-742-923A-6
24	28	77.8	133	4	US-08-533-306A-2
25	28	77.8	133	4	US-08-887-534A-30
26	28	77.8	133	4	US-08-533-306A-6
27	28	77.8	133	4	US-08-742-923A-6
28	28	77.8	133	4	US-08-533-306A-4
29	28	77.8	133	4	US-08-742-923A-6
30	28	77.8	133	4	US-08-533-306A-2
31	28	77.8	133	4	US-08-887-534A-30
32	28	77.8	133	4	US-08-533-306A-6
33	28	77.8	133	4	US-08-742-923A-6
34	28	77.8	133	4	US-08-533-306A-4
35	28	77.8	133	4	US-08-742-923A-6
36	28	77.8	133	4	US-08-533-306A-2
37	28	77.8	133	4	US-08-887-534A-30
38	28	77.8	133	4	US-08-533-306A-6
39	28	77.8	133	4	US-08-742-923A-6
40	28	77.8	133	4	US-08-533-306A-4
41	28	77.8	133	4	US-08-742-923A-6
42	28	77.8	133	4	US-08-533-306A-2
43	28	77.8	133	4	US-08-887-534A-30
44	28	77.8	133	4	US-08-533-306A-6
45	28	77.8	133	4	US-08-742-923A-6

ALIGNMENTS

28	26	72.2	52	4	US-09-588-751-13	Sequence 13, Appl
29	26	72.2	55	4	US-09-015-030-1	Sequence 1, Appl
30	26	72.2	58	4	US-08-819-872A-2	Sequence 2, Appl
31	26	72.2	128	3	US-09-045-764A-6	Sequence 6, Appl
32	26	72.2	158	3	US-08-618-911-4	Sequence 4, Appl
33	26	72.2	158	3	US-08-938-675A-2	Sequence 2, Appl
34	26	72.2	170	1	US-08-127-499A-20	Sequence 20, Appl
35	26	72.2	170	1	US-08-482-847-20	Sequence 20, Appl
36	26	72.2	158	4	US-08-965-056-77	Sequence 77, Appl
37	26	72.2	158	4	US-08-965-056-78	Sequence 78, Appl
38	26	72.2	198	4	US-08-965-056-79	Sequence 80, Appl
39	26	72.2	198	4	US-08-965-056-80	Sequence 81, Appl
40	26	72.2	198	4	US-08-965-056-81	Sequence 81, Appl
41	26	72.2	209	2	US-08-993-228-6	Sequence 6, Appl
42	26	72.2	224	4	US-09-134-001C-3648	Sequence 3648, Ap
43	26	72.2	234	2	US-08-903-801-3	Sequence 3, Appl
44	26	72.2	234	4	US-09-295-055-3	Sequence 3, Appl
45	26	72.2	241	4	US-09-134-001C-5598	Sequence 5598, Ap

RESULT 1
US-08-726-306A-183
Sequence 183, Application US/08726306A
Patent No. 5958684
GENERAL INFORMATION:
APPLICANT: van Leeuwen, Frederik Willem
APPLICANT: Burbach, Johannes Peter Henrl
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1 Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 Inch, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726.306A
FILING DATE: 02-Oct-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080.4
FILING DATE: 02-Oct-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/009,832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 183:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-726-306A-183
Query Match 100.0%; Score 36; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
1111111
DB 121 ELEDKON 127

RESULT 2

US-08-618-911-6
Sequence 6, Application US/08618911
Patent No. 5850016
GENERAL INFORMATION:
APPLICANT: Jung, Rudolf
APPLICANT: Hastings, Craig
APPLICANT: Coughlan, Sean
TITLE OF INVENTION: ALTERATION OF AMINO ACID COMPOSITIONS IN
TITLE OF INVENTION: SEEDS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: 700 Capital Square, 400 Locust Street
CITY: Des Moines
STATE: Iowa
COUNTRY: USA
ZIP: 50309

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,911
FILING DATE: Concurrently herewith
ATTORNEY/AGENT INFORMATION:
NAME: Simon, Soma
REGISTRATION NUMBER: 37,444
REFERENCE/DOCKET NUMBER: 365-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-08-618-911-6
MOLECULE TYPE: protein

Query Match 80.6%; Score 29; DB 2; Length 158;
Best Local Similarity 71.4%; Pred. No. 1.1e+02;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
1111111
DB 124 ELEKEN 130

RESULT 3

US-09-134-001C-4601
Sequence 4601, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674

SEQ ID NO 4601
LENGTH: 196
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4601

Query Match 80.6%; Score 29; DB 4; Length 196;
Best Local Similarity 85.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDKON 7
1111111
DB 59 ELEDKON 65

RESULT 4

US-08-742-753-2
Sequence 2, Application US/08742753
Patent No. 5861278
GENERAL INFORMATION:
APPLICANT: WONG, Gordon G.
APPLICANT: YAO, Kwok-Ming
TITLE OF INVENTION: HNF3-delta Compositions
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,753
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: LAZAR, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5277
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8260
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-08-742-753-2
MOLECULE TYPE: protein

Query Match 80.6%; Score 29; DB 2; Length 771;
Best Local Similarity 71.4%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
1111111
DB 202 ELEKEN 208

RESULT 5

US-08-966-318-5
Sequence 5, Application US/08966318
Patent No. 6001593
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Yue, Henry
APPLICANT: Corley, Neil
APPLICANT: Shah, Purvi

```

; TITLE OF INVENTION: HUMAN MYOSIN HEAVY CHAIN-LIKE PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/966,318
; FILING DATE: Filed Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0413 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532473
; US-08-966-318-5

Query Match
Best Local Similarity 77.8%; Score 28; DB 3; Length 133;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
DB 99 ELDRKEN 105

RESULT 6
US-09-216-619-5
; Sequence 5, Application US/09216619
; Patent No. 6197512
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil
; APPLICANT: Shah, Purni
; TITLE OF INVENTION: HUMAN MYOSIN HEAVY CHAIN-LIKE PROTEINS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/216,619
```

```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/966,318
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0413 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532473
; US-09-216-619-5

Query Match
Best Local Similarity 77.8%; Score 28; DB 4; Length 133;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
DB 99 ELDRKEN 105

RESULT 7
US-08-887-534A-2
; Sequence 2, Application US/08887534A
; Patent No. 6455323
; GENERAL INFORMATION:
; APPLICANT: Holden, David W.
; TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,534A
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Rijn-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 28341/33996
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELE: (312) 474-6600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-887-534A-2

Query Match 77.8%; Score 28; DB 4; Length 419;
```

Best Local Similarity 71.4%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDKON 7
DB 256 ELDNDKEN 262

RESULT 8

US-08-887-534A-7

Sequence 7, Application US/08887534A

Patent No. 6455323

GENERAL INFORMATION:

APPLICANT: Holden, David W.

TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS

NUMBER OF SEQUENCES: 106

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 233 South Wacker Drive/6300 Sears Tower

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/887,534A

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Rin-Laures, Li-Hsien

REGISTRATION NUMBER: 33,547

REFERENCE/DOCKET NUMBER: 28341/33996

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 474-6300

TELEFAX: (312) 474-0448

TELEX: (312) 474-6600

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 419 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-887-534A-7

Query Match 77.8%; Score 28; DB 4; Length 419;
Best Local Similarity 71.4%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDKON 7

DB 256 ELDNDKEN 262

RESULT 9

US-08-533-306A-2

Sequence 2, Application US/08533306A

Patent No. 5837457

GENERAL INFORMATION:

APPLICANT: Liu, Pu

APPLICANT: Collins, Francis S.

APPLICANT: Siciliano, Michael J.

APPLICANT: Claxton, David

TITLE OF INVENTION: Markers for Detection of Chromosome 16

REARRANGEMENTS

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Harnes, Dickey & Pierce, P.L.C.

STREET: P.O. Box 828

CITY: Bloomfield Hills

STATE: MI
COUNTRY: USA
ZIP: 48303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/533,306A

FILING DATE: September 25, 1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Smith, Deann F.

REGISTRATION NUMBER: 36683

REFERENCE/DOCKET NUMBER: 2115-00869COB

TELECOMMUNICATION INFORMATION:

TELEPHONE: (810) 641-1600

TELEFAX: (810) 641-0270

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 576 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-533-306A-2

Query Match 77.8%; Score 28; DB 2; Length 576;
Best Local Similarity 71.4%; Pred. No. 5.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7

DB 242 ELEDKON 248

RESULT 10

US-08-742-923A-2

Sequence 2, Application US/08742923A

Patent No. 5869611

GENERAL INFORMATION:

APPLICANT: Liu, Pu

APPLICANT: Collins, Francis S.

APPLICANT: Siciliano, Michael J.

APPLICANT: Claxton, David

TITLE OF INVENTION: Markers for Detection of Chromosome 16

REARRANGEMENTS

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Harnes, Dickey & Pierce, P.L.C.

STREET: P.O. Box 828

CITY: Bloomfield Hills

STATE: MI

COUNTRY: USA

ZIP: 48303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/742,923A

FILING DATE: No. 5869611ember 1, 1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Smith, Deann F.

REGISTRATION NUMBER: 36683

REFERENCE/DOCKET NUMBER: 2115-00869DVC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (810) 641-1600

TELEFAX: (810) 641-0270

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 576 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-923A-2

Query Match 77.8%; Score 28; DB 2; Length 576;
Best Local Similarity 71.4%; Pred. No. 5.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
DB 242 ELEDERN 248

RESULT 11
US-08-887-534A-30
Sequence 30, Application US/08887534A
Patent No. 6453323

GENERAL INFORMATION:
APPLICANT: Holden, David W.
TITLE OF INVENTION: ANTI-BACTERIAL METHODS AND MATERIALS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/887,534A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Rln-Laureg, Ll-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 28341/33996
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: (312) 474-6600
INFORMATION FOR SEQ. ID NO. 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 729 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-887-534A-30

Query Match 77.8%; Score 28; DB 4; Length 729;
Best Local Similarity 71.4%; Pred. No. 7.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
DB 141 ELDKON 147

RESULT 12
US-08-533-306A-6
Sequence 6, Application US/08533306A
Patent No. 5837457

GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David

TITLE OF INVENTION: Markers for Detection of Chromosome 16
TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/533,306A
FILING DATE: September 25, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-008690B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ. ID NO. 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-533-306A-6

Query Match 77.8%; Score 28; DB 2; Length 816;
Best Local Similarity 71.4%; Pred. No. 8.4e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
DB 482 ELEDERN 488

RESULT 13
US-08-742-923A-6
Sequence 6, Application US/08742923A
Patent No. 5869611

GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,923A
FILING DATE: No. 5869611ember 1, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.

REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869DVC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-923A-6

Query Match 77.8%; Score 28; DB 2; Length 816;
Best Local Similarity 71.4%; Pred. No. 9.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
||||:|
DB 482 ELEDERN 488

RESULT 14
US-08-533-306A-4
Sequence 4, Application US/08533306A
Patent No. 5837457
GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/533,306A
FILING DATE: September 25, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869COB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-533-306A-4

Query Match 77.8%; Score 28; DB 2; Length 885;
Best Local Similarity 71.4%; Pred. No. 9.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
||||:|
DB 551 ELEDERN 557

RESULT 15
US-08-742-923A-4
Sequence 4, Application US/08742923A
Patent No. 5869611
GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,923A
FILING DATE: No. 5869611ember 1, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869DVC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-923A-4

Query Match 77.8%; Score 28; DB 2; Length 885;
Best Local Similarity 71.4%; Pred. No. 9.1e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
||||:|
DB 551 ELEDERN 557

Search completed: November 13, 2002, 13:18:07
Job time : 11.1277 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2002, 13:18:18 : Search time 4.76596 Seconds
(without alignments)
22.121 Million cell updates/sec

Title: US-09-856-086-4
Perfect score: 36
Sequence: 1 ELEDKON 7

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 97044 seqs, 15060890 residues

Total number of hits satisfying chosen parameters: 97044

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	80.6	189	US-09-815-242-5318	Sequence 5318, Ap
2	29	80.6	189	US-09-815-242-12409	Sequence 12409, A
3	29	80.6	189	US-09-815-242-12752	Sequence 12752, A
4	29	80.6	189	US-09-815-242-13142	Sequence 13142, A
5	28	77.8	117	US-09-764-877-1113	Sequence 1113, Ap
6	28	77.8	117	US-09-864-761-42114	Sequence 42114, A
7	28	77.8	211	US-09-864-761-18919	Sequence 48919, A
8	28	77.8	272	US-09-923-300-1697	Sequence 1697, Ap
9	28	77.8	419	US-09-815-242-5273	Sequence 5273, Ap
10	28	77.8	419	US-09-815-242-12154	Sequence 12154, A
11	28	77.8	419	US-09-933-474-2	Sequence 2, Appl
12	28	77.8	529	US-09-888-615-62	Sequence 62, Appl
13	28	77.8	647	US-09-815-242-11218	Sequence 11218, A
14	28	77.8	756	US-09-992-586-62	Sequence 62, Appl
15	28	77.8	756	US-09-989-722-62	Sequence 62, Appl
16	28	77.8	756	US-09-989-723-62	Sequence 62, Appl
17	28	77.8	756	US-09-989-279-62	Sequence 62, Appl
18	28	77.8	756	US-09-989-727-62	Sequence 62, Appl
19	28	77.8	756	US-09-989-731-62	Sequence 62, Appl

20	28	77.8	756	US-09-989-732-62	Sequence 62, Appl
21	28	77.8	756	US-09-991-073-62	Sequence 62, Appl
22	28	77.8	756	US-09-990-442-62	Sequence 62, Appl
23	28	77.8	756	US-09-991-163-62	Sequence 62, Appl
24	28	77.8	756	US-09-993-604-62	Sequence 62, Appl
25	28	77.8	756	US-09-990-456-62	Sequence 62, Appl
26	28	77.8	756	US-09-989-721-62	Sequence 62, Appl
27	28	77.8	848	US-09-925-302-766	Sequence 766, App
28	28	77.8	1085	US-10-084-700-4	Sequence 4, Appl
29	28	77.8	1085	US-09-095-881-2	Sequence 2, Appl
30	28	77.8	2783	US-09-816-669A-14	Sequence 14, Appl
31	27	75.0	41	US-09-864-761-38875	Sequence 38875, A
32	27	75.0	42	US-09-864-761-44621	Sequence 44621, A
33	27	75.0	64	US-09-864-761-46549	Sequence 46549, A
34	27	75.0	78	US-09-864-761-48959	Sequence 48959, A
35	27	75.0	93	US-09-939-980-381	Sequence 381, App
36	27	75.0	223	US-09-925-300-1006	Sequence 1006, Ap
37	27	75.0	331	US-09-815-242-5912	Sequence 5912, Ap
38	27	75.0	331	US-09-815-242-12872	Sequence 12872, A
39	27	75.0	331	US-09-815-242-13146	Sequence 13146, A
40	27	75.0	476	US-09-938-803-12	Sequence 12, Appl
41	27	75.0	489	US-09-876-889-350	Sequence 350, App
42	27	75.0	506	US-09-738-973-34	Sequence 34, Appl
43	27	75.0	576	US-09-794-589-2	Sequence 2, Appl
44	27	75.0	640	US-09-918-951-4	Sequence 4, Appl
45	27	75.0	764	US-09-900-518A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-815-242-5318
Sequence 5318, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA, Olla
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5318
LENGTH: 189
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5318
Query Match 80.6%; Score 29; DB 10; Length 189;
Best Local Similarity 85.7%; Pred. No. 27;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ELEDKON 7
|||||
Db 52 ELEDLON 58

RESULT 2
US-09-815-242-12409
; Sequence 12409, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12409
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12409

Query Match 80.6%; Score 29; DB 10; Length 189;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
|||||
Db 52 ELEDLON 58

RESULT 3
US-09-815-242-12752
; Sequence 12752, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12752
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12752

Query Match 80.6%; Score 29; DB 10; Length 189;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
|||||
Db 52 ELEDLON 58

RESULT 4
US-09-815-242-13142
; Sequence 13142, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13142
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13142

Query Match 80.6%; Score 29; DB 10; Length 189;

Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ELEDKON 7
|||||
Db 52 ELEDLON 58

RESULT 5
US-09-764-877-1113

Sequence 1113, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1113
LENGTH: 41
TYPE: PRT
ORGANISM: Homo sapiens
US-09-764-877-1113

Query Match 77.8%; Score 28; DB 10; Length 41;
Best Local Similarity 83.3%; Pred. No. 8.3;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 2 LEDKON 7
|:|||||
Db 16 LODKON 21

RESULT 6
US-09-864-761-42114

Sequence 42114, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 42114
LENGTH: 117
TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
OTHER INFORMATION: MAP TO AC007845.12
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
OTHER INFORMATION: EST_HUMAN HIT: AW500988.1, EVALU 9.00e-14
OTHER INFORMATION: SWISSPROT HIT: O42184, EVALU 4.50e+00
US-09-864-761-42114

Query Match 77.8%; Score 28; DB 10; Length 117;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ELEDKON 7
|||||
Db 10 ELEDROS 16

RESULT 7
US-09-864-761-48919

Sequence 48919, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecm1ca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/508,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
SEQ ID NO: 48919
LENGTH: 211
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004943.1
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
OTHER INFORMATION: SWISSPROT HIT: P28167, EVALUE 8.00e-22
OTHER INFORMATION: EST_HUMAN HIT: AM66388.1, EVALUE 4.00e-98
US-09-864-761-48919

Query Match
Best Local Similarity 77.8%; Score 28; DB 10; Length 211;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
DB 175 DLEDKOS 181

RESULT 8
US-09-925-300-1697
Sequence 1697, Application US/09925300
Patent No. US20020151681A1
GENERAL INFORMATION:
APPLICANT: Craig Rosen,
APPLICANT: Steve Ruben,
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA101
CURRENT APPLICATION NUMBER: US/09/925,300
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05988
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 1890
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 1697
LENGTH: 272
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (256)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (258)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (262)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (263)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE
LOCATION: (267)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1697

Query Match
Best Local Similarity 77.8%; Score 28; DB 10; Length 272;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
DB 193 EAEEDKEN 199

RESULT 9
US-09-815-242-5273
Sequence 5273, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zykkind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes In
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 5273
LENGTH: 419
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5273

Query Match
Best Local Similarity 77.8%; Score 28; DB 10; Length 419;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
DB 256 ELDKDN 262

RESULT 10
US-09-815-242-12154
Sequence 12154, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zykkind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.

```
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12154
LENGTH: 419
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12154
```

```
Query Match
Best Local Similarity 77.8%; Score 28; DB 10; Length 419;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 ELEDKON 7
DB 256 ELDNKEN 262
```

```
RESULT 11
US-09-932-474-2
Sequence 2, Application US/09924474
Patent No. US20020072105A1
GENERAL INFORMATION:
APPLICANT: BENSON, Timothy
APPLICANT: PRINCE, Donald Bryan
TITLE OF INVENTION: CRYSTALLIZATION AND STRUCTURE DETERMINATION OF FEMA AND FEMA-LIKE
FILE REFERENCE: 00236.US1
CURRENT APPLICATION NUMBER: US/09/932,474
CURRENT FILING DATE: 2001-08-17
PRIOR APPLICATION NUMBER: US 60/226,239
PRIOR FILING DATE: 2000-08-17
PRIOR APPLICATION NUMBER: US 60/226,269
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 419
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-932-474-2
```

```
Query Match
Best Local Similarity 77.8%; Score 28; DB 10; Length 419;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 ELEDKON 7
DB 256 ELDNKEN 262
```

RESULT 12

```
US-09-888-615-62
Sequence 62, Application US/0988615
Patent No. US20020064856A1
GENERAL INFORMATION:
APPLICANT: PLOWMAN, GREGORY
APPLICANT: WHYTE, DAVID
APPLICANT: CAENEPEEL, SEAN
APPLICANT: CHARYDCZAK, GLEN
APPLICANT: MANNING, GERARD
APPLICANT: SUDARSANAM, SUCHA
TITLE OF INVENTION: NOVEL PROTEASES
FILE REFERENCE: 038602/1214
CURRENT APPLICATION NUMBER: US/09/888,615
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: 60/214,047
PRIOR FILING DATE: 2000-06-26
NUMBER OF SEQ ID NOS: 150
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 62
LENGTH: 529
TYPE: PRT
ORGANISM: Homo sapiens
US-09-888-615-62
```

```
Query Match
Best Local Similarity 77.8%; Score 28; DB 10; Length 529;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 ELEDKON 7
DB 238 EMDRON 244
```

```
RESULT 13
US-09-815-242-11218
Sequence 11218, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11218
LENGTH: 647
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11218
```

Query Match 77.8%; Score 28; DB 10; Length 647;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
Db 635 ELEKKN 641

RESULT 14
US-09-992-598-62
Sequence 62, Application US/09992598
Patent No. US20020160384A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlisen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC20
CURRENT APPLICATION NUMBER: US/09/992,598
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952

;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090246
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090252
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090234
;; PRIOR FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090355
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090431
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090435
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090444
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090535
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090540
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090542
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090676
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090678
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090690
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090694
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090696
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090862
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091478
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091544
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091626
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091633
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091978
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/092182
;; PRIOR FILING DATE: 1998-07-09

Query Match

Best Local Similarity 77.8%; Score 28; DB 9; Length 756;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDRON 7

DB 465 ELEDRON 471

RESULT 15
US-09-989-722-62

Sequence 62, Application US/09989722
Patent No. US20020072067A1

GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi J.
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Bolstein, David
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Eaton, Dan L.
;; APPLICANT: Ferrara, Napoleone
;; APPLICANT: Fong, Sherman
;; APPLICANT: Gerber, Hanspeter
;; APPLICANT: Gerltsen, Mary E.
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Kijavir, Ivar J.
;; APPLICANT: Napier, Mary A.
;; APPLICANT: Pan, James
;; APPLICANT: Paoni, Nicholas F.
;; APPLICANT: Roy, Margaret Ann
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Williams, P. Mickey
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;; TITLE OF INVENTION: Acids Encoding the Same
;; FILE REFERENCE: P2730PIC63
;; CURRENT FILING DATE: US/09/989,722
;; PRIOR FILING DATE: 2001-11-19
;; PRIOR APPLICATION NUMBER: 60/049787
;; PRIOR FILING DATE: 1997-06-16
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065186
;; PRIOR FILING DATE: 1997-11-12
;; PRIOR APPLICATION NUMBER: 60/065311
;; PRIOR FILING DATE: 1997-11-13
;; PRIOR APPLICATION NUMBER: 60/066770
;; PRIOR FILING DATE: 1997-11-24
;; PRIOR APPLICATION NUMBER: 60/075945
;; PRIOR FILING DATE: 1998-02-25
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087607
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088021
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029

PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22

PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 77.88; Score 28; DB 10; Length 756;
Best Local Similarity 71.48; Pred. NO. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Caps 0;

Oy 1 ELEDKON 7
Db 465 ELEDKON 471

Search completed: November 13, 2002, 13:40:18
Job time : 4.90881 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:16:18 ; Search time 140.447 seconds
(without alignments)
32.134 Million cell updates/sec

Title: US-09-856-086-4

Perfect score: 36

Sequence: 1 ELEDKCN 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Pending_Patents_AU_Main:*

- 1: /cgn2_6/ptodata/1/paa/US06.COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06.COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07.COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US08.COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US09.COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US09.COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US09.COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US09.COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US09.COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US09.COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US09.COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US09.COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US09.COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US09.COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US10.COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US10.COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US10.COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US10.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	141	22	US-09-834-366-15179
2	36	100.0	141	27	US-60-197-873-15179
3	36	100.0	282	21	US-09-791-537-14762
4	36	100.0	542	21	US-09-791-537-92908
5	36	100.0	542	21	US-09-791-537-125497
6	36	100.0	542	26	US-10-205-331-57

7	36	100.0	543	20	US-09-611-526-3132	Sequence 3132, Ap
8	36	100.0	543	21	US-09-791-537-34475	Sequence 34475, A
9	36	100.0	543	21	US-09-791-537-56903	Sequence 56903, A
10	36	100.0	544	21	US-09-791-537-86802	Sequence 86802, A
11	36	100.0	548	21	US-09-791-537-125495	Sequence 125495, A
12	36	100.0	554	21	US-09-791-537-94465	Sequence 94465, A
13	36	100.0	555	21	US-09-791-537-18635	Sequence 18635, A
14	34	94.4	1175	19	US-09-540-2098-9066	Sequence 9066, Ap
15	33	91.7	368	25	US-10-155-881-11952	Sequence 11952, Ap
16	32	88.9	250	21	US-09-708-427-9494	Sequence 9494, Ap
17	32	88.9	287	21	US-09-708-427-9493	Sequence 9493, Ap
18	32	88.9	357	21	US-09-708-427-9492	Sequence 9492, Ap
19	32	88.9	587	19	US-09-513-996A-52879	Sequence 52879, A
20	32	88.9	624	19	US-09-513-996A-52877	Sequence 52877, A
21	32	88.9	654	19	US-09-513-996A-52876	Sequence 52876, A
22	32	88.9	1170	21	US-09-791-537-24949	Sequence 24949, A
23	32	88.9	1170	27	US-60-360-039-3006	Sequence 3006, Ap
24	32	88.9	1418	21	US-09-791-537-64811	Sequence 64811, A
25	32	88.9	1418	27	US-60-360-039-5052	Sequence 5052, Ap
26	31	86.1	180	20	US-09-614-150-32301	Sequence 32301, A
27	31	86.1	180	27	US-60-173-464-22035	Sequence 22035, A
28	31	86.1	180	27	US-60-191-637-31873	Sequence 31873, A
29	31	86.1	180	27	US-60-191-681-25261	Sequence 25261, A
30	31	86.1	434	27	US-60-337-358-631	Sequence 631, App
31	31	86.1	434	27	US-60-360-039-22371	Sequence 22371, A
32	31	86.1	544	27	US-60-360-039-11132	Sequence 11132, A
33	31	86.1	544	1	PCT-US99-22853B-830	Sequence 830, App
34	31	86.1	590	1	PCT-US99-22853B-829	Sequence 829, App
35	31	86.1	590	18	US-09-451-320-2640	Sequence 2640, Ap
36	31	86.1	603	21	US-09-708-427-25819	Sequence 25819, A
37	31	86.1	668	21	US-09-708-427-25818	Sequence 25818, A
38	31	86.1	674	21	US-09-708-427-25817	Sequence 25817, A
39	31	86.1	895	21	US-09-791-537-96505	Sequence 96505, A
40	31	86.1	895	27	US-60-360-039-1036	Sequence 1036, Ap
41	31	86.1	970	16	US-09-248-796-16191	Sequence 16191, A
42	31	86.1	970	27	US-60-096-409-16191	Sequence 16191, A
43	31	86.1	1091	25	US-10-179-131-8833	Sequence 8833, Ap
44	31	86.1	1091	21	US-09-791-537-98141	Sequence 98141, A
45	31	86.1	1091	27	US-60-360-039-6328	Sequence 6328, Ap

ALIGNMENTS

RESULT 1
US-09-834-366-15179
Sequence 15179, Application US/09834366
GENERAL INFORMATION:
APPLICANT: Benjamin, Stephanie
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Milne Edwards, Jean Baptiste
APPLICANT: Robert, Severin
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81, US2, REG
CURRENT FILING DATE: 2001-04-13
PRIORITY FILING DATE: 2001-04-13
PRIORITY APPLICATION NUMBER: US 60/197,873
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent, pm
SEQ ID NO 15179
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSTRE
LOCATION: 21
OTHER INFORMATION: Xaa - Gly, Arg
US-09-834-366-15179
Query Match 100.0%; Score 36; DB 22; Length 141;
Best Local Similarity 100.0%; Pred. No. 40;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
|||||||

Db 57 ELEDKON 63

RESULT 2
US-60-197-873-15179
Sequence 15179, Application US/60197873
GENERAL INFORMATION:
APPLICANT: Beljantin, Stephane
APPLICANT: Tanaka, Hiroaki
APPLICANT: Dumas Maline Edwards, Jean Baptiste
APPLICANT: Jobert, Severin
APPLICANT: Giordano, Jean-Yves
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: 81 US1.PRO
CURRENT APPLICATION NUMBER: US/60/197,873
CURRENT FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 52153
SOFTWARE: Patent.pm
SEQ ID NO 15179
LENGTH: 141
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: UNSURE
LOCATION: 21
OTHER INFORMATION: Xaa = Gly, Arg
US-60-197-873-15179

Query Match
Best Local Similarity 100.0%; Score 36; DB 27; Length 141;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
|||||||

Db 57 ELEDKON 63

RESULT 3
US-09-791-537-14762
Sequence 14762, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patent version 3.0
SEQ ID NO 14762
LENGTH: 282
TYPE: PRT
ORGANISM: Mus musculus
US-09-791-537-14762

Query Match
Best Local Similarity 100.0%; Score 36; DB 21; Length 282;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
|||||||

Db 95 ELEDKON 101

RESULT 4
US-09-791-537-92908
Sequence 92908, Application US/09791537

GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patent version 3.0
SEQ ID NO 92908
LENGTH: 542
TYPE: PRT
ORGANISM: Mus musculus
US-09-791-537-92908

Query Match
Best Local Similarity 100.0%; Score 36; DB 21; Length 542;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
|||||||

Db 335 ELEDKON 341

RESULT 5
US-09-791-537-125497
Sequence 125497, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patent version 3.0
SEQ ID NO 125497
LENGTH: 542
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-791-537-125497

Query Match
Best Local Similarity 100.0%; Score 36; DB 21; Length 542;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
|||||||

Db 336 ELEDKON 342

RESULT 6
US-10-205-331-57
Sequence 57, Application US/10205331
GENERAL INFORMATION:
APPLICANT: Warner-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Alistair
APPLICANT: Brooksbank, Robert
APPLICANT: Pluncock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WL-A-018199
CURRENT APPLICATION NUMBER: US/10/205,331
CURRENT FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: GB 0118354.0
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 117
SOFTWARE: Patent ver. 2.1
SEQ ID NO 57

LENGTH: 542
TYPE: PRT
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: NF-L
US-10-205-331-57

Query Match 100.0%; Score 36; DB 26; Length 542;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
Db 336 ELEDKON 342

RESULT 7
US-09-611-526-3132
Sequence 3132, Application US/09611526
GENERAL INFORMATION:
APPLICANT: OTA, TOSHIO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: ISOGAI, TAKAO
APPLICANT: HAYASHI, KOJI
APPLICANT: ISHII, SHIZUKO
APPLICANT: KAMAI, YURI
APPLICANT: WAKAMATSU, AI
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: NAGAI, KEIICHI
APPLICANT: KOJIMA, SHINICHI
APPLICANT: OHSUKI, TETSUJI
APPLICANT: KOGA, HISASHI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIS OF FULL LENGTH CDNAS
FILE REFERENCE: 08335/0122
CURRENT APPLICATION NUMBER: US/09/611,526
CURRENT FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: JP 1999-194486
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: JP 2000-118774
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183765
PRIOR FILING DATE: 2000-05-02
NUMBER OF SEQ ID NOS: 4484
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3132
LENGTH: 543
TYPE: PRT
ORGANISM: Homo sapiens
US-09-611-526-3132

Query Match 100.0%; Score 36; DB 20; Length 543;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
Db 335 ELEDKON 341

RESULT 8
US-09-791-537-34475
Sequence 34475, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0
SEQ ID NO 34475
LENGTH: 543
TYPE: PRT
ORGANISM: Mus musculus
US-09-791-537-34475

Query Match 100.0%; Score 36; DB 21; Length 543;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
Db 336 ELEDKON 342

RESULT 9
US-09-791-537-56903
Sequence 56903, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 56903
LENGTH: 543
TYPE: PRT
ORGANISM: Mus musculus
US-09-791-537-56903

Query Match 100.0%; Score 36; DB 21; Length 543;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
Db 336 ELEDKON 342

RESULT 10
US-09-791-537-86802
Sequence 86802, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 86802
LENGTH: 544
TYPE: PRT
ORGANISM: Homo sapiens
US-09-791-537-86802

Query Match 100.0%; Score 36; DB 21; Length 544;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
Db 336 ELEDKON 342

```
RESULT 11
US-09-791-537-125495
; Sequence 125495, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent version 3.0
; LENGTH: 548
; SEQ ID NO 125495
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-791-537-125495
```

```
Query Match          100.0%; Score 36; DB 21; Length 548;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ELEDKON 7
Db 334 ELEDKON 340
```

```
RESULT 12
US-09-791-537-94465
; Sequence 94465, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent version 3.0
; SEQ ID NO 94465
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-791-537-94465
```

```
Query Match          100.0%; Score 36; DB 21; Length 554;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ELEDKON 7
Db 335 ELEDKON 341
```

```
RESULT 13
US-09-791-537-18635
; Sequence 18635, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent version 3.0
```

```
; SEQ ID NO 18635
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-791-537-18635
```

```
Query Match          100.0%; Score 36; DB 21; Length 555;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ELEDKON 7
Db 336 ELEDKON 342
```

```
RESULT 14
US-09-540-209B-9066
; Sequence 9066, Application US/09540209B
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDE
; FILE REFERENCE: 2709,1001-001
; CURRENT APPLICATION NUMBER: US/09/540,209B
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 9066
; LENGTH: 1175
; TYPE: PRT
; ORGANISM: B.fragilis
US-09-540-209B-9066
```

```
Query Match          94.4%; Score 34; DB 19; Length 1175;
Best Local Similarity 85.7%; Pred. No. 1.5e+03;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ELEDKON 7
Db 825 ELEDKON 831
```

```
RESULT 15
US-10-155-881-11952
; Sequence 11952, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyya, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(15300)J
; CURRENT APPLICATION NUMBER: US/10/155,881
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 11952
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Zea mays
US-10-155-881-11952
```

```
Query Match          91.7%; Score 33; DB 25; Length 368;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 ELEDKON 7
Db 276 ELEDKON 282
```

Search completed: November 13, 2002, 13:39:17
Job time : 141.447 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:17:03 : Search time 2.82979 Seconds
(Without alignments)
29.431 Million cell updates/sec

Title: US-09-856-086-4

Perfect score: 36

Sequence: 1 ELEDKON 7

Scoring table: BLOSUM62

Searched: Gapop 10.0, Gapext 0.5

Total number of hits satisfying chosen parameters: 41632

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/pdata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/pdata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/pdata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/pdata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/pdata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/pdata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/pdata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	7	US-09-856-086-4	Sequence 4, Appl
2	31	86.1	103	US-09-513-999C-6211	Sequence 6211, Ap
3	29	80.6	196	US-10-092-411A-4601	Sequence 4601, Ap
4	28	77.8	756	US-10-131-819A-392	Sequence 392, App
5	28	77.8	756	US-10-131-819A-392	Sequence 392, App
6	28	77.8	756	US-10-131-823A-392	Sequence 392, App
7	28	77.8	756	US-10-131-823A-392	Sequence 392, App
8	28	77.8	756	US-10-131-823A-392	Sequence 392, App
9	28	77.8	756	US-10-131-823A-392	Sequence 392, App
10	28	77.8	756	US-10-131-823A-392	Sequence 392, App
11	28	77.8	756	US-10-125-926A-392	Sequence 392, App
12	28	77.8	756	US-10-127-831A-392	Sequence 392, App
13	28	77.8	756	US-10-127-831A-392	Sequence 392, App
14	28	77.8	756	US-10-127-831A-392	Sequence 392, App
15	28	77.8	756	US-10-127-831A-392	Sequence 392, App
16	28	77.8	756	US-10-127-831A-392	Sequence 392, App
17	28	77.8	756	US-10-127-831A-392	Sequence 392, App
18	28	77.8	756	US-10-127-831A-392	Sequence 392, App
19	28	77.8	756	US-10-127-831A-392	Sequence 392, App
20	28	77.8	756	US-10-127-831A-392	Sequence 392, App
21	28	77.8	756	US-10-131-833A-392	Sequence 392, App
22	28	77.8	756	US-10-131-833A-392	Sequence 392, App
23	28	77.8	756	US-10-125-930A-392	Sequence 392, App
24	28	77.8	756	US-10-127-825A-392	Sequence 392, App
25	28	77.8	756	US-10-127-838B-392	Sequence 392, App
26	28	77.8	756	US-10-127-843A-392	Sequence 392, App

27	28	77.8	756	6	US-10-128-684A-392	Sequence 392, App
28	28	77.8	756	6	US-10-128-685A-392	Sequence 392, App
29	28	77.8	756	6	US-10-128-686A-392	Sequence 392, App
30	28	77.8	756	6	US-10-128-690A-392	Sequence 392, App
31	28	77.8	756	6	US-10-128-693A-392	Sequence 392, App
32	28	77.8	756	6	US-10-131-821A-392	Sequence 392, App
33	28	77.8	756	6	US-10-131-836A-392	Sequence 392, App
34	28	77.8	756	6	US-10-137-872A-392	Sequence 392, App
35	28	77.8	756	6	US-10-137-873A-392	Sequence 392, App
36	28	77.8	756	6	US-10-125-921A-392	Sequence 392, App
37	28	77.8	756	6	US-10-125-928A-392	Sequence 392, App
38	28	77.8	756	6	US-10-127-821A-392	Sequence 392, App
39	28	77.8	756	6	US-10-127-822A-392	Sequence 392, App
40	28	77.8	756	6	US-10-127-824A-392	Sequence 392, App
41	28	77.8	756	6	US-10-127-827A-392	Sequence 392, App
42	28	77.8	756	6	US-10-127-830A-392	Sequence 392, App
43	28	77.8	756	6	US-10-127-832A-392	Sequence 392, App
44	28	77.8	756	6	US-10-127-834A-392	Sequence 392, App
45	28	77.8	756	6	US-10-127-836A-392	Sequence 392, App

ALIGNMENTS

```

RESULT 1
US-09-856-086-4
; Sequence 4, Application US/09856086
; GENERAL INFORMATION:
; APPLICANT: EBRINGER, ALAN
; TITLE OF INVENTION: DIAGNOSIS OF DEMYELINATING OR SPONGIFORM DISEASE
; FILE REFERENCE: 78104.040
; CURRENT APPLICATION NUMBER: US/09/856,086
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens, Bos
US-09-856-086-4

Query Match      100.0%; Score 36; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. NO. 3.8e+04;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ELEDKON 7
DB      1 ELEDKON 7

RESULT 2
US-09-513-999C-6211
; Sequence 6211, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59. US. REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6211
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 90
; OTHER INFORMATION: Xaa-Leu or Val

```


US-09-513-999C-6211

Query Match 86.1%; Score 31; DB 5; Length 103;
Best Local Similarity 85.7%; Pred. No. 4.1;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
| | | | |
DB 89 ELEDKON 95

RESULT 3

US-10-092-411A-4601
Sequence 4601, Application US/10092411A
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm at al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: 032796-101
CURRENT APPLICATION NUMBER: US/10/092,411A
CURRENT FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: US 09/134,001
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5676
SEQ ID NO 4601
LENGTH: 196
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-10-092-411A-4601

Query Match 80.6%; Score 29; DB 6; Length 196;
Best Local Similarity 85.7%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
| | | | |
DB 59 ELEDKON 65

RESULT 4

US-10-131-813A-392
Sequence 392, Application US/10131813A
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C139
CURRENT APPLICATION NUMBER: US/10/131,813A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26

PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 392
LENGTH: 756
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-813A-392

Query Match 77.8%; Score 28; DB 6; Length 756;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
| | | | |
DB 465 ELEDKON 471

RESULT 5

US-10-131-819A-392
Sequence 392, Application US/10131819A
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P330R1C134
CURRENT APPLICATION NUMBER: US/10/131,819A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17

SEQ ID NO 392
LENGTH: 756

1 ELEDKÖN 7

Db 465 EAEDRON 471

RESULT 8

US-10-131-826A-392

Sequence 392, Application US/10131826A
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C128
CURRENT APPLICATION NUMBER: US/10/131,826A
CURRENT FILING DATE: 2002-04-24
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 392
LENGTH: 756
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-826A-392

Query Match 77.8%; Score 28; DB 6; Length 756;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDRON 7

Db 465 EAEDRON 471

RESULT 9

US-10-131-829A-392

Sequence 392, Application US/10131829A
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C138
CURRENT APPLICATION NUMBER: US/10/131,829A
CURRENT FILING DATE: 2002-04-27
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 392
LENGTH: 756
TYPE: PRT
ORGANISM: Homo Sapien
US-10-131-829A-392

Query Match 77.8%; Score 28; DB 6; Length 756;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDRON 7

Db 465 EAEDRON 471

RESULT 10

US-10-125-926A-392

Sequence 392, Application US/10125926A
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C80
CURRENT APPLICATION NUMBER: US/10/125, 926A
PRIOR FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 350
SEQ ID NO 392
LENGTH: 756
TYPE: PRT
ORGANISM: Homo Sapien
US-10-125-926A-392

Query Match 77.8%; Score 28; DB 6; Length 756;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
DB 465 ELEDKON 471

RESULT 11
US-10-127-829A-392
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C85
CURRENT APPLICATION NUMBER: US/10/127, 829A
PRIOR FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: 60/049911

PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 350
SEQ ID NO 392
LENGTH: 756
TYPE: PRT
ORGANISM: Homo Sapien
US-10-127-829A-392

Query Match 77.8%; Score 28; DB 6; Length 756;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
DB 465 ELEDKON 471

RESULT 12
US-10-127-831A-392
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C107
CURRENT APPLICATION NUMBER: US/10/127, 831A
PRIOR FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122

```
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/059184
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/059263
;; PRIOR FILING DATE: 1997-09-18
;; PRIOR APPLICATION NUMBER: 60/059352
;; PRIOR FILING DATE: 1997-09-19
;; PRIOR APPLICATION NUMBER: 60/059588
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 392
;; LENGTH: 756
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-127-831A-392
```

```
Query Match      77.8%; Score 28; DB 6; Length 756;
Best Local Similarity 71.4%; Pred. NO. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 ELEDKON 7
DB      465 EAEDRON 471
```

RESULT 13

```
US-10-127-835A-392
; Sequence 392, Application US/10127835A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C102
; CURRENT APPLICATION NUMBER: US/10/127, 835A
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; Remaining Prior Application data removed - See File Wrapper or PALM.
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 392
;; LENGTH: 756
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-127-835A-392
```

```
Query Match      77.8%; Score 28; DB 6; Length 756;
Best Local Similarity 71.4%; Pred. NO. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 ELEDKON 7
DB      465 EAEDRON 471
```

RESULT 14

```
US-10-127-837A-392
; Sequence 392, Application US/10127837A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C96
; CURRENT APPLICATION NUMBER: US/10/127, 837A
; PRIOR FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; Remaining Prior Application data removed - See File Wrapper or PALM.
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 392
;; LENGTH: 756
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-127-837A-392
```

```
Query Match      77.8%; Score 28; DB 6; Length 756;
Best Local Similarity 71.4%; Pred. NO. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      1 ELEDKON 7
DB      465 EAEDRON 471
```

OY 1 ELEDKON 7
DB 465 ELEDKON 471

RESULT 15

US-10-127-842A-392
; Sequence 392. Application US/10127842A

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACTS ENCODING THE SAME
FILE REFERENCE: P330R1C100
CURRENT FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: US/10/127, 842A
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 392
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-127-842A-392

Query Match 77.8%; Score 28; DB 6; Length 756;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDKON 7
DB 465 ELEDKON 471

Search completed: November 13, 2002, 13:39:39
Job time: 3.82979 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using SW model

Run on: November 13, 2002, 13:13:25 ; Search time 10.8723 Seconds

(without alignments)
61.895 Million cell updates/sec

Title: US-09-856-086-4

Sequence: 1 ELEDKON 7

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	36	100.0	287 2 A21762	neurofilament trip
2	36	100.0	543 1 QFMSL	neurofilament trip
3	36	100.0	544 2 S07144	neurofilament trip
4	36	100.0	548 1 QFGL	neurofilament trip
5	36	100.0	554 2 JM0094	neurofilament trip
6	32	88.9	357 2 T06746	neurofilament trip
7	32	88.9	1170 2 A72287	neurofilament trip
8	32	88.9	1418 2 T15332	neurofilament trip
9	31	86.1	434 2 S46605	neurofilament trip
10	31	86.1	544 2 B44841	neurofilament trip
11	31	86.1	655 2 C71438	neurofilament trip
12	31	86.1	674 2 B71438	neurofilament trip
13	31	86.1	887 2 G88484	neurofilament trip
14	31	86.1	895 2 E64431	neurofilament trip
15	31	86.1	1091 2 T34107	neurofilament trip
16	30	83.3	286 2 H90200	neurofilament trip
17	30	83.3	293 2 S15320	neurofilament trip
18	30	83.3	293 2 T43586	neurofilament trip
19	30	83.3	400 2 E83675	neurofilament trip
20	30	83.3	564 2 F50992	neurofilament trip
21	30	83.3	715 2 S77439	neurofilament trip
22	30	83.3	750 2 S73829	neurofilament trip
23	30	83.3	855 2 C72858	neurofilament trip
24	30	83.3	855 2 B89472	neurofilament trip
25	30	83.3	978 2 H81311	neurofilament trip
26	30	83.3	1328 2 T18897	neurofilament trip
27	30	83.3	1818 1 S73852	neurofilament trip
28	30	83.3	2701 2 S17796	neurofilament trip
29	29	80.6	168 2 PS0179	neurofilament trip

30	29	80.6	179 2 S52934	alkyl hydroperoxid
31	29	80.6	189 2 E89804	alkyl hydroperoxid
32	29	80.6	263 2 E70215	hypothetical prote
33	29	80.6	281 2 F83923	manganese-containi
34	29	80.6	302 2 T37514	hypothetical prote
35	29	80.6	358 2 T25010	hypothetical prote
36	29	80.6	486 2 C81332	hypothetical prote
37	29	80.6	523 2 T19684	hypothetical prote
38	29	80.6	534 2 H82244	sensor histidine k
39	29	80.6	573 2 T21355	hypothetical prote
40	29	80.6	582 2 S22195	hypothetical prote
41	29	80.6	617 2 S37744	endo-exonuclease y
42	29	80.6	643 2 A97234	bps2 protein - Des
43	29	80.6	676 2 H87906	ABC-type transport
44	29	80.6	678 2 F71921	protein F23H2.6 [1
45	29	80.6	850 1 T05180	hypothetical prote
				S-receptor kinase

ALIGNMENTS

RESULT 1

A21762 neurofilament triplet L protein - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 27-Jul-1990 #sequence_revision 27-Jul-1990 #text_change 13-Aug-1999

C:Accession: A21762

R:Jullien, J.P., Ramchandran, K., Grosfeld, F.

Biochim. Biophys. Acta 825, 398-404, 1985

A:Title: Cloning of a cDNA encoding the smallest neurofilament protein from the rat.

A:Reference number: A21762; PMID:85252830; PMID:3925999

A:Accession: A21762

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-287 <YUL>

A:Cross-references: GB:M25638; MID:g205683; PIDN:AAA1694.1; PID:g205684

C:Superfamily: cytoskeletal keratin

C:Keywords: coiled coil

Query Match	100.0%	Score 36;	DB 2;	Length 287;
Best Local Similarity	100.0%	Pred. No. 6.2;		
Matches 7;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 ELEDKON 7			
DB	139 ELEDKON 145			

RESULT 2

QFMSL neurofilament triplet L protein - mouse

N:Alternate names: 68k neurofilament protein; NF-L(low) protein; type IV IF protein

C:Species: Mus musculus (house mouse)

C>Date: 31-Mar-1988 #sequence_revision 30-Sep-1993 #text_change 22-Jun-1999

C:Accession: A25227; A26562; A43772; A41012; I55316

R:Lewis, S.A.; Cowan, N.J.

Mol. Cell. Biol. 6, 1529-1534, 1986

A:Title: Anomalous placement of introns in a member of the intermediate filament mult

A:Reference number: A25227; PMID:87064433; PMID:3785173

A:Accession: A25227

A:Molecule type: DNA

A:Residues: 1-543 <LEW>

A:Cross-references: GB:M13016; MID:g200023; PIDN:AAA39810.1; PID:g387492

A>Note: The authors translated the codon GGC for residue 5 as Ala, NCA for residue 88

1 as Glu

J. Cell Biol. 100, 843-850, 1985

A:Title: Genetics, evolution, and expression of the 68,000-mol-wt neurofilament prote

A:Reference number: A26562; PMID:85131334; PMID:3919033

A:Accession: A26562

A:Molecule type: mRNA

A:Residues: 242-543 <LEW>

A:Cross-references: GB:X02165

A:Experimental source: brain
 R:Julien, J.P.; Meyer, D.; Flavell, D.; Hurst, J.; Grosveld, F.
 Brain Res. Mol. Brain Res. 1, 243-250, 1986
 A:Title: Cloning and developmental expression of the murine neurofilament gene family.
 A:Reference number: A43772
 A:Accession: A43772
 A:Molecule type: mRNA
 A:Residues: 1-5, 'Y', '7-8, 'Y', '10-64, 'M', '66-72, 'L', '74-98, 'D', '100-194, 'R', '196-202, 204-239, 'Y'
 A:Cross-references: GB:M20480; NID:9200037; PIDN:AAA39814.1; PID:9200038
 A:Note: the authors translated the codon CCG for residue 195 as Ala
 J: Biol. Chem. 266, 18861-18867, 1991
 A:Title: Identification of Ser-55 as a major protein kinase A phosphorylation site on th
 A:Reference number: A41012; MUID:92011653; PMID:1171455
 A:Accession: A41012
 A:Molecule type: protein
 A:Residues: 52-57 <SIH>
 R:Wakihara, K.; Ikenaka, K.; Wade, K.; Tamura, T.
 J. Biol. Chem. 265, 19786-19791, 1990
 A:Title: Structure of the 68-kDa neurofilament gene and regulation of its expression.
 A:Reference number: I55316; MUID:91060592; PMID:2246261
 A:Accession: I55316
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-5, 'Y', '7-8, 'Y', '10-28 <RES>
 A:Cross-references: GB:M5423; NID:9200027; PIDN:AAA39812.1; PID:9554245
 A:Comment: This is the most abundant of the three neurofilament proteins and, as the oth
 C:Genetics:
 A:introns: 349/3; 391/2; 498/1
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:2-73/Domain: head <HED>
 F:94-125/Domain: coil 1a, alpha-helical rod #status predicted <R1A>
 F:126-138/Domain: linker 1
 F:139-234/Domain: coil 1b, alpha-helical rod #status predicted <R1B>
 F:235-256/Region: linker 12
 F:257-272/Domain: coil 2a, alpha-helical rod #status predicted <R2A>
 F:273-281/Region: linker 2
 F:282-401/Domain: coil 2b, alpha-helical rod #status predicted <R2B>
 F:404-543/Domain: tail <TAI>
 F:404-444/Region: tail subdomain a
 F:445-543/Region: tail subdomain b

Query Match 100.0%; Score 36; DB 1; Length 543;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 |||||
 Db 336 ELEDKON 342

RESULT 3
 507144
 neurofilament triplet L protein - human
 N:Alternate names: neurofilament light polypeptide (68k)
 C:Species: Homo sapiens (man)
 C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 21-Jul-2000
 C:Accession: S07144; I52832; A60703
 R:Julien, J.P.; Grosveld, F.; Yazdanhsh, K.; Flavell, D.; Meljer, D.; Mushynski, W.
 Biochem. Biophys. Acta 909, 10-20, 1997
 A:Title: The structure of a human neurofilament gene (NF-L): a unique exon-intron organ
 A:Reference number: 507144; MUID:87214213; PMID:3034332
 A:Accession: 507144
 A:Molecule type: DNA
 A:Residues: 1-544 <JUL>
 A:Cross-references: EMBL:X05608; NID:91495072; PIDN:CAA29097.1; PID:91279504
 A:Note: the authors translated the codon ATG for residue 366 as Asn
 R:Popelov, V.A.; Popelova, T.V.; Julien, J.P.
 Cell Growth Differ. 5, 187-196, 1994
 A:Title: Ap-1 and Krox-24 transcription factors activate the neurofilament light gene pr
 A:Reference number: I52832; MUID:94235564; PMID:8180132

A:Accession: I52832
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-10 <POS>
 A:Cross-references: GB:S70309; NID:9547176; PIDN:AAD4057.1; PID:94261757
 R:Nomata, Y.; Watanabe, T.; Wade, H.
 J. Biochem. 93, 825-831, 1983
 A:Title: Highly acidic proteins from human brain: purification and properties of GI
 A:Reference number: A60703; MUID:83265667; PMID:6135695
 A:Accession: A60703
 A:Molecule type: protein
 A:Residues: 469-472, 'D', '474 <NOM>
 A:Experimental source: Glu-50 brain peptide
 A:Note: this acidic protein is named for its greater than fifty per cent glutamic a
 C:Genetics:
 A:Gene: GDB:NEFL, NFL
 A:Cross-references: GDB:120227; OMIM:162280
 A:Map position: 8p21-8p21
 A:introns: 349/3; 391/2; 498/1
 C:Superfamily: cytoskeletal keratin
 C:Keywords: brain; coiled coil; intermediate filament
 F:469-544/Product: Glu-50 peptide #status predicted <E50>

Query Match 100.0%; Score 36; DB 2; Length 544;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 |||||
 Db 336 ELEDKON 342

RESULT 4
 507144
 neurofilament triplet L protein - pig
 N:Alternate names: 68k neurofilament protein
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 15-Nov-1984 #sequence_revision 28-May-1986 #text_change 10-Oct-1997
 C:Accession: A01337; A90973; A34569; A02963
 R:Gelsler, N.; Plessmann, U.; Weber, K.
 FEBS Lett. 182, 475-478, 1985
 A:Title: The complete amino acid sequence of the major mammalian neurofilament prot
 A:Reference number: A01337; MUID:85154583; PMID:3920075
 A:Accession: A01337
 A:Molecule type: protein
 A:Residues: 1-547 <GE1>
 R:Gelsler, N.; Kauffmann, E.; Fischer, S.; Plessmann, U.; Weber, K.
 EMBO J. 2, 1295-1302, 1983
 A:Title: Neurofilament architecture combines structural principles of intermediate :
 A:Reference number: A90973
 A:Accession: A90973
 A:Molecule type: protein
 A:Residues: 1-82, 278-548 <GE2>
 A:Note: residue 322 is either lysine or arginine
 R:Gonda, Y.; Nishizawa, K.; Ando, S.; Kitamura, S.; Minoura, Y.; Nishit, Y.; Inagaki,
 Biochem. Biophys. Res. Commun. 167, 1316-1325, 1990
 A:Title: Involvement of protein kinase C in the regulation of assembly-disassembly
 A:Reference number: A34569; MUID:90211318; PMID:2108674
 A:Accession: A34569
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 9-14; 23-29; 30-53 <GON>
 C:Comment: Mammalian neurofilaments usually contain three polypeptides, L, M, and H
 ke all other intermediate filament proteins; a conserved alpha-helical region, whose
 al beta turns; domain b is acidic and rich in glutamic acid and lysine residues.
 C:Comment: The extra mass and high charge density that distinguish the neurofilament
 charged scaffolding structure suitable for interaction with other neuronal component
 C:Comment: The boundaries of the domain between residues 70-92 and 399-402 are not
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament
 F:1-70/Domain: head <HED>

F:92-123/Domain: coll 1a, alpha-helical rod <RIA>
 F:137-232/Domain: coll 1b, alpha-helical rod <RIb>
 F:255-399/Domain: coll 2, alpha-helical rod <RIb>
 F:402-548/Domain: tail <TAI>
 F:402-442/Region: tail subdomain a
 F:443-548/Region: tail subdomain b

Query Match 100.0%; Score 36; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
 |||||
 Db 334 ELEDKON 340

RESULT 5

JM0094
 neurofilament protein-L - bovine

N:Alternate names: NF-L

C:Species: Bos primigenius taurus (cattle)

C>Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 07-May-1999

C:Accession: JM0094

R:Hashimoto, R.; Nakamura, Y.; Goto, H.; Wada, Y.; Sakoda, S.; Kalduchi, K.; Inagaki, M.

Biochem. Biophys. Res. Commun. 245, 407-411, 1998

A:Title: Domain- and site-specific phosphorylation of bovine NF-L by Rho-associated kinase

A:Reference number: JM0094; PMID:98238650; PMID:9571164

A:Accession: JM0094

A:Molecule type: protein

A:Residues: 1-554 <RMS>

A:Comment: Domain- and site-specific phosphorylation by Rho-kinase regulates the assembly

C:Superfamily: cytoskeletal keratin

Query Match 100.0%; Score 36; DB 2; Length 554;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
 |||||
 Db 335 ELEDKON 341

RESULT 6

T06746
 hypothetical protein F15B8.70 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 17-Mar-2000

C:Accession: T06746

R:Queller, F.; Benes, V.; Rechmann, S.; Borkova, D.; Ansorge, W.; Salanoubat, M.; Mewes,

submitted to the Protein Sequence Database, April 1999

A:Reference number: Z15794

A:Accession: T06746

A:Molecule type: DNA

A:Residues: 1-357 <QNE>

A:Cross-references: EMBL:AL049660; GSPDB:GNO0061; ATSP:F15B8.70

A:Experimental source: cultivar Columbia; BAC clone F15B8

C:Genetics:

A:Gene: ATSP:F15B8.70

A:Map position: 3

C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 88.9%; Score 32; DB 2; Length 357;
 Best Local Similarity 85.7%; Pred. No. 52;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7
 |||||
 Db 273 ELEDKON 279

RESULT 7

A72287
 hypothetical protein TM1182 - Thermotoga maritima (strain MS8)

C:Species: Thermotoga maritima
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: A72287
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic
 Garret, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome

A:Reference number: A72200; PMID:99287316; PMID:10360571

A:Accession: A72287

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1170 <ARN>

A:Cross-references: GB:AE001774; GB:AE000512; NID:g4981717; PIDN:AAD36257.1; PID:g498

A:Genetics:

A:Gene: TM1182

C:Superfamily: chromosome segregation protein SMC1

Query Match 88.9%; Score 32; DB 2; Length 1170;
 Best Local Similarity 85.7%; Pred. No. 1,7e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
 |||||
 Db 422 ELEDKON 428

RESULT 8

T15232
 hypothetical protein B0414.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T15232

R:Sammons, L.; Wohldmann, P.; Rohlfing, T.

submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid B0414.

A:Reference number: Z18312

A:Accession: T15232

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1418 <SAR>

A:Cross-references: EMBL:AF003145; NID:g2088760; PID:g2088765; PIDN:AAB57718.1; GSPDB

A:Experimental source: strain Bristol N2; clone B0414

C:Genetics:

A:Gene: CESP:B0414.7

A:Map position: 1

A:introns: 30/2; 61/1; 85/3; 206/3; 282/1; 369/3; 580/3; 842/3; 886/3; 921/3; 1090/3;

Query Match 88.9%; Score 32; DB 2; Length 1418;
 Best Local Similarity 85.7%; Pred. No. 2.1e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
 |||||
 Db 1082 ELEDKON 1088

RESULT 9

S46605
 26S proteasome regulatory particle chain RPTs - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein O3258; protein YOR117W; protein YOR3258W; YTA1 protein

C:Species: Saccharomyces cerevisiae

C>Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 19-Jan-2001

C:Accession: S46605; S60993; S61675; S67002; S63870; S34352

R:Schmell, R.; Mannhaupt, G.; Stucke, R.; Tauer, R.; Ehle, S.; Schwarze, C.; Velt

Yeast 10, 1141-1155, 1994

A:Title: Identification of a set of yeast genes coding for a novel family of putative

A:Reference number: S46605; PMID:95274317; PMID:7754704

A:Accession: S46605

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-434 <SCH>

A:Cross-references: EMBL:X73569; NID:g13877; PIDN:CAA51971.1; PID:g13878
 R:Wiemann, S.; Rechmann, S.; Benes, V.; Voss, H.; Schwager, C.; Vlcek, C.; Stegemann, J.
 Submitted to the EMBL Data Library, August 1995
 A:Description: Sequencing of 51 kilobases on the right arm of chromosome XV from *S. cere*
 A:Reference number: 560983
 A:Accession: 560993

A:Molecule type: DNA
 A:Residues: 1-434 <WIE>

A:Cross-references: EMBL:X90518; NID:g1050808; PIDN:CAA62114.1; PID:g1050819
 R:Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Banrevel, A.; Sander, C.; Valencle

submitted to the EMBL Data Library, December 1995
 A:Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome

A:Reference number: 561643
 A:Accession: 561675

A:Molecule type: DNA
 A:Residues: 1-434 <BEN>

A:Cross-references: EMBL:X94335; NID:g1262139; PIDN:CAA64037.1; PID:g1164962
 R:Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansoorge, W.

submitted to the Protein Sequence Database, July 1996
 A:Reference number: 56965
 A:Accession: 567002

A:Molecule type: DNA
 A:Residues: 1-434 <VOS>

A:Cross-references: EMBL:X75025; NID:g1420310; PIDN:CAA99315.1; PID:g1420311; MIPS:YOR11
 A:Experimental source: strain S288C

R:Wiemann, S.; Rechmann, S.; Benes, V.; Voss, H.; Schwager, C.; Vlcek, C.; Stegemann, J.
 Yeast 12, 281-288, 1996

A:Title: Sequencing and analysis of 51 kb on the right arm of chromosome XV from *Sacchar*
 A:Reference number: 563860; MUID:97060020; PMID:8904341

A:Accession: 563870
 A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-434 <WIM>

A:Cross-references: EMBL:X90518; NID:g1050808; PIDN:CAA62114.1; PID:g1050819
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995

C:Genetics:
 A:Gene: SCD:RPT5; YTA1

A:Cross-references: SCD:S0005643; MIPS:YOR117W
 A:Map position: 15R

C:Superfamily: ATP-dependent 26S protease: FtsH/SEC18/CDC48-type ATP-binding domain hc

C:Keywords: ATP, nucleotide binding, P-loop
 F:193-403/Domain: FtsH/SEC18/CDC48-type ATP-binding domain motif A (P-loop)

F:222-229/Region: nucleotide-binding motif A (P-loop)

Query Match
 Best Local Similarity 86.1%; Score 31; DB 2; Length 434;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 I:|||||

Db 90 EIEDKEN 96

RESULT 10

low molecular weight neurofilament protein XNF-L - African clawed frog
 C:Species: *Xenopus laevis* (African clawed frog)
 C:Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995

C:Accession: B44841
 R:Charms, L.R.; Szabo, B.G.; Garner, H.

J. Neurosci. 13, 3010-3024, 1992
 A:Title: Identification and developmental expression of a novel low molecular weight neu

A:Reference number: A44841; MUID:92356194; PMID:1494944
 A:Accession: B44841

A:Molecule type: mRNA
 A:Status: preliminary

A:Residues: 1-544 <CHA>
 A:Experimental source: brain

A:Note: sequence inconsistent with the nucleotide translation
 C:Superfamily: cytoskeletal keratin

Query Match
 86.1%; Score 31; DB 2; Length 544;

Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 I:|||||

Db 329 EIEDKOS 335

RESULT 11

hypothetical protein - Arabidopsis thaliana
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 A:Variety: Columbia

C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
 R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.;
 P.; Wedler, H.; Wedler, E.; Wambutt, R.; Wellenreger, T.; Pohl, T.M.; Terry, N.;
 Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulgdom
 erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechmann, S.;
 C.; Chalwatizis, N.
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis*
 A:Reference number: A71400; MUID:98121113; PMID:9461215
 A:Accession: B71438

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-674 <BEV>
 A:Cross-references: GB:597342; NID:g2245031; PID:g2245057
 C:Genetics:
 A:Map position: 4COP9-4G3845

Query Match
 Best Local Similarity 86.1%; Score 31; DB 2; Length 674;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 I:|||||

Db 412 EIEDKEN 418

RESULT 13

G88484

protein F23F12.8 [Imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: G88484

R:Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; NID:99069613; PMID:9651916

A:Note: see webistes genome.wustl.edu/96/C_elegans/ and www.sanger.ac.uk/Projects/C_elg

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: G88484

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-887 <STO>

A:Cross-References: GB:chr_III; PIDN:AAA20610.1; PID:9529217; GSPDB:GN00021; CESP:F23F12

A:Note: exon 5 similar to trichohyalin and tropomyosin

C:Genetics:

A:Gene: F23F12.8

A:Map position: 3

Query Match

Best Local Similarity 86.1%; Score 31; DB 2; Length 887;

Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7

DB 544 ELEDKON 550

RESULT 14

E64431

UDPglucose 6-dehydrogenase (EC 1.1.1.22) (intra-cellular) - Methanococcus jannaschii

N:Contains: Intein

C:Species: Methanococcus jannaschii

C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: E64431

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,

Reich, C.I.; Overbeek, R.; Kikines, E.F.; Weiss, K.G.; Merrick, J.M.; Glodek, A.;

Reon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A:Reference number: A64300; NID:96337999; PMID:8688087

A:Accession: E64431

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-895 <BUL>

A:Cross-References: GB:U67548; GB:L77117; NID:92826361; PIDN:AA99056.1; PID:91591706; T

C:Genetics:

A:Map position: FOR990967-993654

C:Keywords: oxidoreductase; protein splicing

Query Match

Best Local Similarity 86.1%; Score 31; DB 2; Length 895;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7

DB 286 ELEDKON 292

RESULT 15

T34107

hypothetical protein C18C4.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T34107

R:Gatung, S.

submitted to the EMBL Data Library, April 1996

A:Description: The sequence of C. elegans cosmid C18C4.

A:Reference number: Z21478

A:Accession: T34107

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1091 <GAT>

A:Cross-References: EMBL:U55369; PIDN:AA23825.1; GSPDB:GN00023; CESP:C18C4.5

A:Experimental source: strain Bristol N2; clone C18C4

C:Genetics:

A:Gene: CESP:C18C4.5

A:Map position: 5 38/3; 95/3; 179/2; 259/3; 301/3; 419/2; 573/3; 613/1; 875/3; 920/2; 959/2;

A:Introns: 38/3; 95/3; 179/2; 259/3; 301/3; 419/2; 573/3; 613/1; 875/3; 920/2; 959/2;

Query Match

Best Local Similarity 86.1%; Score 31; DB 2; Length 1091;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ELEDKON 7

DB 660 ELEDKON 666

Search completed: November 13, 2002, 13:23:28
Job time : 13.0152 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 13, 2002, 13:13:24 ; Search time 5.80851 Seconds
(without alignments)
49.984 Million cell updates/sec

Title: US-09-856-086-4
Perfect score: 36
Sequence: 1 ELEDKON 7

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	541	1 NFL_RAT	P19527 rattus norv
2	36	100.0	542	1 NFL_MOUSE	P08551 mus musculu
3	36	100.0	543	1 NFL_HUMAN	P07196 homo sapien
4	36	100.0	548	1 NFL_PIG	P02547 sus scrofa
5	36	100.0	554	1 NFL_BOVIN	P02548 bos taurus
6	31	86.1	434	1 PRSA_YEAST	P33297 saccharomyc
7	31	86.1	544	1 NFL_XENLA	P35616 xenopus lae
8	31	86.1	887	1 YLX8_CAEEL	P45604 caenorhabdi
9	31	86.1	885	1 YAS4_MERUA	P58454 methanococc
10	30	83.3	252	1 YOPN_YERPE	P16160 yerania pe
11	30	83.3	750	1 YD33_MYCPN	P75445 mycoplasma
12	30	83.3	808	1 Y066_NPVAC	P41467 autographa
13	30	83.3	1281	1 IP35_MOUSE	P09239 mus musculu
14	30	83.3	1818	1 HMY2_MYCPN	P75471 mycoplasma
15	30	83.3	2701	1 IP35_HUMAN	P14571 homo sapien
16	30	83.3	2701	1 IP35_HUMAN	P29995 rattus norv
17	29	80.6	168	1 YNDH_BACSP	P26830 bacillus sp
18	29	80.6	302	1 YAO3_SCHPO	P01082 schizosacch
19	29	80.6	519	1 TRPE_SERMA	P00897 serratia ma
20	29	80.6	582	1 BP52_ACTAM	P33285 actinomyces
21	29	80.6	617	1 RNC1_YEAST	P33753 saccharomyc
22	29	80.6	759	1 FXM1_RAT	P97691 rattus norv
23	29	80.6	760	1 FXM1_MOUSE	P08696 mus musculu
24	29	80.6	859	1 MUTS_AQUAE	O08652 aquifex aeo
25	29	80.6	3674	1 SPGR_HUMAN	O09652 aquifex aeo
26	29	80.6	4967	1 RYR2_HUMAN	O09236 homo sapien
27	29	80.6	4969	1 RYR2_HUMAN	O09236 homo sapien
28	28	77.8	104	1 SCPL_BACSU	P81099 bacillus su
29	28	77.8	228	1 Y516_BORBU	O51668 borrelia bu
30	28	77.8	281	1 Y4RD_RHISN	P55637 rhizobium su
31	28	77.8	333	1 Y591_AQUAE	O62653 aquifex aeo
32	28	77.8	361	1 DBPA_RAT	O62164 rattus norv
33	28	77.8	365	1 PURK_AQUAE	O66608 aquifex aeo

ALIGNMENTS

RESULT 1	ID	NFL_RAT	STANDARD	PRT	541 AA.
AC	P19527	063367			
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Neurofilament triplet L protein (68 kDa neurofilament protein)				
DE	(Neurofilament light polypeptide) (NF-L).				
GN	NFL OR NFL OR NF68.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90184052; PubMed=2516804;				
RA	Chin S.S., Liem R.K.H.;				
RT	"Expression of rat neurofilament proteins NF-L and NF-M in				
RT	transfected non-neuronal cells."				
RL	Eur. J. Cell Biol. 50:475-490(1989).				
RN	[2]				
RP	SEQUENCE OF 197-483 FROM N.A.				
RX	MEDLINE=95252830; PubMed=3925999;				
RA	Julien J.-P., Ramachandran K., Grosfeld F.;				
RT	"Cloning of a cDNA encoding the smallest neurofilament protein from				
RT	the rat."				
RL	Biochim. Biophys. Acta 825:398-404(1985).				
RN	[3]				
RP	SEQUENCE OF 1-10 FROM N.A.				
RX	MEDLINE=95264348; PubMed=7745611;				
RA	Reeben M., Neuman T., Palgi J., Palm K., Paalme V., Saarma M.;				
RT	"Characterization of the rat light neurofilament (NF-L) gene promoter				
RT	and identification of NGF and cAMP responsive regions."				
RL	J. Neurosci. Res. 40:177-186(1995).				
RN	[4]				
RP	CARBOHYDRATE-LINKAGE SITES.				
RX	MEDLINE=93346421; PubMed=8344946;				
RA	Dong D.L.-Y., Xu Z.-S., Chevrier M.R., Cotter R.J., Cleveland D.W.,				
RT	Hart G.W.;				
RT	"Glycosylation of mammalian neurofilaments. Localization of multiple				
RT	O-linked N-acetylglucosamine moieties on neurofilament polypeptides				
RL	L and M."				
RL	J. Biol. Chem. 268:16679-16687(1993).				
CC	-1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,				
CC	AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.				
CC	-1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH				
CC	THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT				
CC	A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH				
CC	OTHER NEURONAL COMPONENTS OR IONS.				
CC	-1- PTM: O-GLYCOSYLATED; CONTAINS THREE N-ACETYLGLUCOSAMINE SIDE				
CC	CHAINS.				
CC	-1- MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE				
CC	NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL				
CC	INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM				


```

FT SITE 381 391 EPI TOPE (RECOGNIZED BY IF-SPECIFIC
FT CONFLICT 5 5 MONOCLONAL ANTIBODY.
FT CONFLICT 8 8 Y -> S (IN REF. 1).
FT CONFLICT 64 64 Y -> I (IN REF. 1).
FT CONFLICT 72 72 M -> K (IN REF. 1).
FT CONFLICT 98 98 D -> H (IN REF. 1).
FT CONFLICT 194 194 R -> A (IN REF. 1).
FT CONFLICT 202 202 MISSING (IN REF. 2).
FT CONFLICT 239 239 Y -> I (IN REF. 1).
SQ SEQUENCE 542 AA; 61448 MW; 8EE9B8C6F0831D8C CRC64;

Query Match 100.0%; Score 36; DB 1; Length 542;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
DB 335 ELEDKON 341

RESULT 3
ID NFL_HUMAN STANDARD; PRT; 543 AA.
AC P07196; Q16154;
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament triplet polypeptide) (NF-L).
GN NEFL OR NEFL OR NF68.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87214213; PubMed=3034332;
RA Julien J.-P., Grosveld F., Yazdankhsh K., Flavell D., Meijer D.,
RA Mushynski W.,
RT "The structure of a human neurofilament gene (NF-L): a unique exon-
RT intron organization in the intermediate filament gene family.",
RL Biochim. Biophys. Acta 909:10-20(1987).
RN [2]
RP SEQUENCE OF 1-10 FROM N.A.
RA MEDLINE=9435564; PubMed=8180132;
RA Pospelov V.A., Pospelova T.V., Julien J.-P.,
RA Ap-1 and Krox-24 transcription factors activate the neurofilament
RT light gene promoter in p19 embryonal carcinoma cells.",
RL Cell Growth Differ. 5:187-196(1994).
RN [3]
RP VARIANT CMT2E PRO-332.
RA MEDLINE=20307176; PubMed=10841809;
RA Meriyanova I.V., Perepelov A.V., Polyakov A.V., Sitnikov V.F.,
RA Dadelis E.L., Oparin R.B., Petrini A.N., Evgarfov O.V.,
RA "A new variant of Charcot-Marie-Tooth disease type 2 is probably the
RT result of a mutation in the neurofilament-light gene.",
RL Am. J. Hum. Genet. 67:37-46(2000).
RN [4]
RP FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPEACE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC -1- DISEASE: DEFECTS IN NEFL ARE A CAUSE OF CHARCOT-MARIE-TOOTH
CC DISEASE TYPE 2E (CMT2E).
CC -1- MISCELLANEOUS: NEFL IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -1- DATABASE: NAME=Inherited peripheral neuropathies mutation db;

```

```

CC WWW="http://molgen-www.uia.ac.be/CMTMutations/".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X05608; CAA29097.1; -
DR EMBL: S70309; AAD14057.1; -
DR EMBL: S07144; S07144.
DR Genew; HGNC: 7739; NEFL.
DR MIM; 162280; -
DR InterPro; IPR001664; IF.
DR Pfam; PF00038; filament; 1.
DR PROSITE; PS00236; IF; 1.
KM Intermediate filament; Coiled coil; Neurone; Glycoprotein;
FT Disease mutation; Charcot-Marie-Tooth disease.
FT INIT_MET 0
FT DOMAIN 1 91 HEAD.
FT DOMAIN 92 396 ROD.
FT DOMAIN 397 543 TAIL.
FT DOMAIN 92 123 COIL 1A.
FT DOMAIN 124 136 LINKER 1.
FT DOMAIN 137 234 COIL 1B.
FT DOMAIN 235 252 LINKER 12.
FT DOMAIN 253 271 COIL 2A.
FT DOMAIN 272 280 LINKER 2.
FT DOMAIN 281 396 COIL 2B.
FT DOMAIN 397 443 TAIL, SUBDOMAIN A.
FT DOMAIN 444 543 TAIL, SUBDOMAIN B (ACIDIC).
FT CARBOHYD 20 26 O-LINKED (GICNAC) (BY SIMILARITY).
FT CARBOHYD 26 26 O-LINKED (GICNAC) (BY SIMILARITY).
FT SITE 381 391 EPI TOPE (RECOGNIZED BY IF-SPECIFIC
FT VARIANT 332 332 MONOCLONAL ANTIBODY.
FT FT O -> P (IN CMT2E).
SQ SEQUENCE 543 AA; 61645 MW; 7A0F1ADD5BED22F6 CRC64;

Query Match 100.0%; Score 36; DB 1; Length 543;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ELEDKON 7
DB 335 ELEDKON 341

RESULT 4
ID NFL_PIG STANDARD; PRT; 548 AA.
AC P02547;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurofilament triplet L protein (68 kDa neurofilament protein)
DE (Neurofilament light polypeptide) (NF-L).
GN NEFL.
OS Sus scrofa (Pig)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RA MEDLINE=85154583; PubMed=3920075;
RA Geisler N., Plessmann U., Weber K.,
RA "The complete amino acid sequence of the major mammalian
RT neurofilament protein (NF-L).",
RL FEBS Lett. 182:475-478(1985).
RN [2]
RP SEQUENCE OF 1-82 AND 278-548.

```

```

RA Geisler N., Kaufmann E., Fischer S., Plesmann U., Weber K.;  

RT "Neurofilament architecture combines structural principles of  

intermediate filaments with carboxy-terminal extensions increasing  

in size between triplet proteins."; RL EMBO J. 2:1995-1302(1983).  

CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,  

AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.  

CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH  

THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT  

PROTEINS ARE DUE TO THE TAILPIECE EXTENSIONS. THIS REGION MAY FORM  

A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH  

OTHER NEURONAL COMPONENTS OR IONS.  

CC MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE  

NEUROFILAMENT PROTEINS AND, AS THE OTHER NONBETHEIDIAL  

INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM  

FILAMENTS.  

CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.  

DR PR; A02963; QEPGL.  

DR InterPro; IPR001664; IF.  

DR Pfam; PF00038; filament; 1.  

DR ProSITE; PS00226; IF; 1.  

KW Intermediate filament; Coiled coil; Neuron; Glycoprotein.  

FT DOMAIN 1  

FT DOMAIN 92 395  

FT DOMAIN 396 548  

FT DOMAIN 92 123  

FT DOMAIN 124 136  

FT DOMAIN 137 232  

FT DOMAIN 233 251  

FT DOMAIN 252 270  

FT DOMAIN 271 279  

FT DOMAIN 280 395  

FT DOMAIN 396 442  

FT DOMAIN 443 548  

FT CARBOHYD 20 20  

FT CARBOHYD 26 26  

FT SITE 380 390  

FT UNSURE 322 322  

SQ SEQUENCE 548 AA; 61940 MW; 83044813637AC739 CRC64;  

OR K.  

Query Match 100.0%; Score 36; DB 1; Length 548;  

Best Local Similarity 100.0%; Pred. No. 4.8;  

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  

QY 1 ELEDKRON 7  

Db 334 ELEDKRON 340  

|||||  

RESULT 5  

NFL_BOVIN STANDARD; PRT; 554 AA.  

ID NFL_BOVIN P02548; P79127;  

DT 21-JUL-1986 (Rel. 01, Created)  

DT 01-NOV-1997 (Rel. 35, Last sequence update)  

DT 16-OCT-2001 (Rel. 40, Last annotation update)  

DE Neurofilament triplet L protein (68 kDa neurofilament protein)  

DE (Neurofilament light polypeptide) (NF-L) (Micro glutamic acid-rich  

protein).  

DE NEFL.  

GN NEFL.  

OS Bos taurus (Bovine).  

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  

OC Bovidae; Bovinae; Bos.  

CX NCBI_TaxID=9913;  

RN [1]  

RP SEQUENCE FROM N.A.  

RP STRAIN-Holstein; TISSUE-Brain;  

RA Hill W.D., Zhang L., Ball B.J., Spinkley T.J.;  

RT Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  

[2]  

RP SEQUENCE OF 468-554.

```

```

RA MEDLINE-85154567; PubMed-3884373;
RX Iasobe T., Okuyama T.;
RT "Brain micro glutamic acid-rich protein is the C-terminal endpiece of
the neurofilament 68-kDa protein as determined by the primary
RT sequence.";
RL FEBS Lett. 182:389-392(1985).
CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
CC -1- DOMAIN: THE EXTRA MASS AND HIGH CHARGE DENSITY THAT DISTINGUISH
CC THE NEUROFILAMENT PROTEINS FROM ALL OTHER INTERMEDIATE FILAMENT
CC PROTEINS ARE DUE TO THE TAILPEACE EXTENSIONS. THIS REGION MAY FORM
CC A CHARGED SCAFFOLDING STRUCTURE SUITABLE FOR INTERACTION WITH
CC OTHER NEURONAL COMPONENTS OR IONS.
CC MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
CC NEUROFILAMENT PROTEINS AND AS THE OTHER NONEPITHELIAL
CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
CC FILAMENTS.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL: D83919; AAB41543.1; -.
CC DR PIR: A02964; OPRO.
CC DR InterPro: IPR001664; IF.
CC DR Pfam: PF00038; Filament; 1.
CC DR PROSITE: PS00226; IF; 1.
CC KW Intermediate filament; Coiled coil; Neutrone.
CC FT INIT MET 0 0
CC FT DOMAIN 1 92 BY SIMILARITY.
CC FT DOMAIN 93 396 HEAD (BY SIMILARITY).
CC FT DOMAIN 397 554 ROD (BY SIMILARITY).
CC FT DOMAIN 554 554 TAIL (BY SIMILARITY).
CC FT DOMAIN 93 124 COIL 1A.
CC FT DOMAIN 125 137 COIL 1A.
CC FT DOMAIN 138 233 LINDER 1.
CC FT DOMAIN 234 252 COIL 1B.
CC FT DOMAIN 253 271 COIL 1B.
CC FT DOMAIN 272 280 COIL 2A.
CC FT DOMAIN 281 396 COIL 2A.
CC FT DOMAIN 397 443 LINDER 2.
CC FT DOMAIN 443 554 TAIL, SUBDOMAIN A.
CC FT DOMAIN 494 500 TAIL, SUBDOMAIN B (ACIDIC).
CC FT CONFLICT 494 500 MISSING (IN REF. 2).
CC FT CONFLICT 509 509 A -> ABA (IN REF. 2).
CC SQ SEQUENCE 554 AA; 62514 MW; D77281CA2C1CIA CRC64;

Query Match 100.0%; Score 36; DB 1; Length 554;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

02 1 ELEDKQN 7
|||||
Db 335 ELEDKQN 341

RESULT 6
PRSA_YEAST PRSA_YEAST STANDARD; PRT; 434 AA.
AC P33297;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 265 protease regulatory subunit 6A (Tat-binding protein homolog 1)
DE (TBP-1).
GN RPT5 OR YTA1 OR YOR17M OR Q3256 OR YOR3256M.
OS Saccharomyces cerevisiae (Baker's yeast).
CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI_TaxID=4932;
CX

```

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C836;
 RA MEDLINE-95274317; PubMed-754704;
 RA Schnall R., Manhaupt G., Stucke R., Tauer R., Ehmlé S.,
 RA Schwarze C., Vetter I., Feldmann H.,
 RT "Identification of a set of yeast genes coding for a novel family of
 RT putative ATPases with high similarity to constituents of the 26S
 RT protease complex."
 RL Yeast 10:1141-1155(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-8288C / FY1679;
 RA MEDLINE-97060020; PubMed-8904341;
 RA Wemmann S., Rechmann S., Benes V., Voss H., Schwager C., Vlcek C.,
 RA Siegmund J., Zimmermann J., Erfle H., Paces V., Ansgore W.,
 RT "Sequencing and analysis of 51 kb on the right arm of chromosome XV
 RT from Saccharomyces cerevisiae reveals 30 open reading frames."
 RL Yeast 12:281-288(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-97344368; PubMed-9200815;
 RA Voss H., Benes V., Andrae M.A., Valencio A., Rechmann S., Teodoru C.,
 RA Schwager C., Paces V., Sander C., Ansgore W.,
 RT "DNA sequencing and analysis of 130 kb from yeast chromosome XV,"
 RL Yeast 13:653-672(1997).
 CC -1- FUNCTION: THE 26S PROTEASE IS INVOLVED IN THE ATP-DEPENDENT
 CC DEGRADATION OF UBQUITINATED PROTEINS. THE REGULATORY (OR ATPASE)
 CC COMPLEX CONFERS ATP DEPENDENCY AND SUBSTRATE SPECIFICITY TO THE
 CC 26S COMPLEX (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC and nuclear (Potential).
 CC -1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X73569; CAA51971.1; -
 DR EMBL: X80518; CAA62114.1; -
 DR EMBL: X84335; CAA64037.1; -
 DR EMBL: Z75025; CAA93315.1; -
 DR PIR: S46605; S46605.
 DR SGD: S0005643; RPTS.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003959; AAA_ATPase_centre.
 DR InterPro: IPR003960; AAA_SBD.
 DR Pfam: PF00004; AAA_1.
 DR SMART: SM00382; AAA_1.
 DR TIGRFAMs: TIGR01242; 26S945; 1.
 DR PROSITE: PS00674; AAA_1.
 KW Proteasome; ATP-binding; Nuclear protein.
 FT NP_BIND 222 229 ATP (POTENTIAL).
 FT SEQUENCE 434 AA; 48255 MW; C215422033FCE830 CRC64;
 QY Query Match 86.1%; Score 31; DB 1; Length 434;
 DB Best Local Similarity 71.4%; Pred. No. 40;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ELEDKON 7
 DB 90 EIEDKEN 96
 RESULT 7
 ID NPL_XENLA STANDARD; PRT; 544 AA.
 AC P35616;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurofilament triplet L protein (Neurofilament light polypeptide) (NF-
 DE L).
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Metachela; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RA MEDLINE-92356194; PubMed-1494944;
 RA Charnas L.R., Szaro B.G., Gainer H.,
 RT "Identification and developmental expression of a novel low molecular
 RT weight neuronal intermediate filament protein expressed in Xenopus
 RT laevis."
 RL J. Neurosci. 12:3010-3024(1992).
 CC -1- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,
 CC AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.
 CC MISCELLANEOUS: NF-L IS THE MOST ABUNDANT OF THE THREE
 CC NEUROFILAMENT PROTEINS AND, AS THE OTHER NONEPITHELIAL
 CC INTERMEDIATE FILAMENT PROTEINS, IT CAN FORM HOMOPOLYMERIC 10-NM
 CC FILAMENTS.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M86654; AA83018.1; -
 DR PIR: B44841; B44841.
 DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; Filament; 1.
 DR PROSITE: PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Neuron.
 FT DOMAIN 1 87
 FT DOMAIN 88 87 HEAD.
 FT DOMAIN 88 390 ROD.
 FT DOMAIN 391 544 TAIL.
 FT DOMAIN 88 119 COIL_1A.
 FT DOMAIN 120 132 LINKER_1.
 FT DOMAIN 133 228 COIL_1B.
 FT DOMAIN 229 246 LINKER_12.
 FT DOMAIN 247 265 COIL_2A.
 FT DOMAIN 266 274 COIL_2B.
 FT DOMAIN 275 390 COIL_2B.
 FT DOMAIN 391 435 TAIL, SUBDOMAIN A.
 FT DOMAIN 436 544 TAIL, SUBDOMAIN B (ACIDIC).
 FT DOMAIN 441 538 GLU-RICH.
 FT DOMAIN 464 469 POLY-GLU.
 FT SEQUENCE 544 AA; 61861 MW; 76D911B896E97201 CRC64;
 QY Query Match 86.1%; Score 31; DB 1; Length 544;
 DB Best Local Similarity 85.7%; Pred. No. 51;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ELEDKON 7
 DB 329 ELEDKOS 335
 RESULT 8
 ID YLX8_CAEEL STANDARD; PRT; 887 AA.
 AC P46504;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE Hypothetical 104.7 kDa protein F23F12.8 in chromosome III precursor.
 GN F23F12.8.

OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 RX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Du 2;
 RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U12965; AAA20610.1;
 DR WormPep: F23F12.8; CE01255.
 DR InterPro: IPR002557; Chitin_bind_Pera.
 DR SMART: SM00494; ChtBD2; 2.
 DR Hypothetical protein: Signal.
 FT SIGNAL. 1 20
 FT CHAIN 1 887
 FT CARBOHYD 187
 FT CARBOHYD 190
 FT SEQUENCE 887 AA; 104705 MW; 214AC6CA3A1FAL CRC64;
 SO
 Query Match 86.1%; Score 31; DB 1; Length 887;
 Best Local Similarity 71.4%; Pred. No. 85;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ELEDKON 7
 DB 544 EMDRON 550
 RESULT 9
 ID VAS4_METJA STANDARD; PRT; 895 AA.
 AC 058454;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MJ1054 (EC 1.1.1.-) [contains: Mja UDPGD
 DE Intein].
 GN MJ1054.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 RX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 RC MEDLINE=96337999; PubMed=8688087;
 RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Butt G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Welch J.F., Fuhrman J.L., Nguyen D.,
 RA Outterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Huret M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RA jannaschii".
 RT Science 273:1058-1073 (1996).
 RL -1- PTV: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
 CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
 CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE UDP-GLUCOSE/GDP-MANNOSE DEHYDROGENASES
 CC FAMILY.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: D67548; AAB99056.1;
 DR TIGR: MJ1054;
 DR InterPro: IPR003586; Hedgehog_hntc.
 DR InterPro: IPR003587; Hedgehog_hntn.
 DR InterPro: IPR002203; Intein.
 DR InterPro: IPR004042; Intein_endonuc.
 DR InterPro: IPR000205; NAD_binding.
 DR InterPro: IPR01732; UDPG_MGDP_dh.
 DR Pfam: PF03720; UDPG_MGDP_dh_1.
 DR Pfam: PF03721; UDPG_MGDP_dh_N_1.
 DR SMART: SM00305; Hntc_1.
 DR SMART: SM00306; Hntc_1.
 DR PROSITE: PS50818; INTEIN_C_TER; 1.
 DR PROSITE: PS50819; INTEIN_ENDONUCLEASE; 1.
 DR PROSITE: PS50817; INTEIN_N_TER; 1.
 DR Hypothetical protein: Oxidoreductase; NAD; Autocatalytic cleavage;
 DR protein splicing; Complete proteome.
 FT CHAIN 1 260
 FT CHAIN 1 261
 FT CHAIN 715
 FT NP_BIND 2 19
 FT ACT_SITE 261
 FT SEQUENCE 895 AA; 102614 MW; B511A68CCA1445 CRC64;
 SO
 Query Match 86.1%; Score 31; DB 1; Length 895;
 Best Local Similarity 85.7%; Pred. No. 86;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ELEDKON 7
 DB 286 ELEDKON 292
 RESULT 10
 ID YOPN_YERPE STANDARD; PRT; 292 AA.
 AC P16160;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Outer membrane protein yopN (YopN) (LcrE).
 GN YOPN OR LCR E OR YPCD1.39C OR Y0042.
 OS Yersinia pestis, and
 OS Yersinia pseudotuberculosis.
 OC Plasmid pCD1, and Plasmid PIB1.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 RX NCBI_TaxID=632, 633;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Y.pestis; STRAIN-KIM5; PLASMID-PCD1;
 RC MEDLINE=96427122; PubMed=9746557;
 RA Perry R.D., Straley S.C., Fetherston J.D., Rose D.J., Gregor J.,
 RA Blattner F.R.;
 RA "DNA sequencing and analysis of the low-Ca²⁺-response plasmid pCD1 of
 RA Yersinia pestis KIM5".
 RT Infect. Immun. 66:4611-4623 (1998).
 RL [2]
 CC SEQUENCE FROM N.A.
 CC SPECIES-Y.pestis; STRAIN-KIM5; PLASMID-PCD1;
 CC MEDLINE=96427122; PubMed=9746557;
 CC Hu P., Elliott J., McCreedy P., Skowronski E., Garnea J.,
 CC

RA Kobayashi A., Brubaker R.R., Garcia E.;
 RT "Structural organization of virulence-associated plasmids of Yersinia
 pestis";
 RL J. Bacteriol. 180:5192-5202(1998).
 RN (3)
 RP SEQUENCE FROM N.A.
 RC SPECIES-Y-pestis; STRAIN-CO-92 / Bloyer Orientalis; PLASMID-PCOL;
 RX MEDLINE-21470413; PubMed-11586360;
 RA Parthill J., Wren B.W., Thomson N.R., Tiltball R.W., Holden M.T.G.,
 RA Partridge M.B., Sebaldia M., James K.D., Churche C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks R.M., Davis P., Dougan G.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Letwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Simmonds M., Skelton J., Oyston P.C.F., Quail M., Rutherford K.,
 RA "Genome sequence of Yersinia pestis, the causative agent of plague";
 RL Nature 413:523-527(2001).
 RN (4)
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-8.
 RC SPECIES-Y-pseudotuberculosis; STRAIN-YPR11; PLASMID-P1B1;
 RX MEDLINE-91312137; PubMed-1857212;
 RA Forsberg A., Viltanen A.-M., Skurnik M., Wolf-Watz H.;
 RT "The surface-located YopN protein is involved in calcium signal
 transduction in Yersinia pseudotuberculosis";
 RL Mol. Microbiol. 5:977-986(1991).
 CC -1- FUNCTION: PLAYS A MAJOR ROLE IN REGULATION OF THE LOW-CALCIUM
 CC RESPONSE. SEEMS TO SENSE THE CALCIUM CONCENTRATION AND TO TRANSMIT
 CC A SIGNAL TO SHUT OFF YOP TRANSCRIPTION WHEN THE CALCIUM
 CC CONCENTRATION IS HIGH.
 CC -1- SUBCELLULAR LOCATION: Outer membrane.
 CC -1- SIMILARITY: TO Y. ENTEROCOLITICA YOPN.
 CC -1- SIMILARITY: TO S. TYPHIMORIUM INVE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF074612; AAC69792.1; -;
 DR EMBL; AF053946; AAC62565.1; -;
 DR EMBL; AL117189; CAB54916.1; -;
 DR EMBL; X51833; CA36129.1; -;
 DR PIR; S08669; S08669;
 DR PIR; S15320; S15320;
 KW Outer membrane; Calcium; Plasmid; Virulence; Complete proteome.
 FT INIT_MET 0
 FT SEQUENCE 292 AA; 32538 MW; 873DAA041E70877E CRC64;
 SO
 Query Match 83.3%; Score 30; DB 1; Length 292;
 Best Local Similarity 85.7%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ELEDKON 7
 DB 95 ELEDKON 101
 RESULT 11
 YD3_MYCPN
 ID YD3_MYCPN STANDARD; PRT; 750 AA.
 AC P75445;
 DT 01-NOV-1997 (Rel. 35 Created)
 DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Hypothetical protein MPN333 (P10_Orf750).
 GN MPN333 OR MP503.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN (1)

RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 29342 / M129;
 RX MEDLINE-97105885; PubMed-6948633;
 RA Himmelfreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 pneumoniae";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: TO M. PNEUMONIAE MPN335.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000049; AAB96151.1; -;
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 2 22
 FT TRANSMEM 33 53
 FT TRANSMEM 79 99
 FT TRANSMEM 116 136
 FT TRANSMEM 143 163
 FT TRANSMEM 724 744
 FT SEQUENCE 750 AA; 85312 MW; 5CF64FDE600E5A CRC64;
 SO
 Query Match 83.3%; Score 30; DB 1; Length 750;
 Best Local Similarity 71.4%; Pred. No. 11e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ELEDKON 7
 DB 565 ELEDKON 571
 RESULT 12
 Y066_NPVAC
 ID Y066_NPVAC STANDARD; PRT; 808 AA.
 AC P41467;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32; Last sequence update)
 DT 01-NOV-1997 (Rel. 35; Last annotation update)
 DE Hypothetical 94.0 kDa protein in POL-TEB3 intergenic region.
 OS Autographa californica nuclear polyhedrosis virus (ACNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 CC Nucleopolyhedrovirus.
 OX NCBI_TaxID=46015;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-C6; PubMed-8030224;
 RX MEDLINE-94303173; PubMed-8030224;
 RA Ayres M.D., Howard S.C., Kuzio J., Loper-Ferber M., Possee R.D.;
 RT "The complete DNA sequence of Autographa californica nuclear
 RT polyhedrosis virus";
 RL Virology 202:586-605(1994).
 CC -1- SIMILARITY: TO CORRESPONDING ORF IN OPNPV AND LDMPV.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L22858; AAA6696.1; -;
 KW Hypothetical protein.
 FT DOMAIN 106 117
 FT SEQUENCE 808 AA; 93973 MW; 76A871D2B6633F8A CRC64;
 RN (1)

Query Match 83.3%; Score 30; DB 1; Length 808;
 Best Local Similarity 71.4%; Pred. No. 1.2e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 DB 452 QLEDKON 458

RESULT 13
 IP3S_MOUSE
 ID IP3S_MOUSE STANDARD; PRT: 1281 AA.
 AC Q92329; P70226; O61744;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Inositol 1,4,5-trisphosphate receptor type 2 (Type 2 inositol 1,4,5-
 triphosphate receptor) (Type 2 InsP3 receptor) (IP3 receptor isoform
 2) (InsP3R2) (inositol 1,4,5-trisphosphate type V receptor)
 DE (Fragment).
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE OF 1-272 FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE-Heart;
 RX MEDLINE-98399819; PubMed-9729462;
 RA Futatsugi A., Kuwajima G., Mikoshiba K.;
 RT "Muscle-specific mRNA isoform encodes a protein composed mainly of the
 N-terminal 175 residues of type 2 Ins(1,4,5)P3 receptor.";
 RL Biochem. J. 334:559-563(1998).
 RN [2]
 RP SEQUENCE OF 273-1281 FROM N.A.
 RC STRAIN-C3H; TISSUE-Embryo;
 RX MEDLINE-97218118; PubMed-9065779;
 RA De Smedt H., Misiashen L., Parys J.B., Henning R.H., Stenaert I.,
 RA Vanlinden S., Gijssels A., Himpe B., Casteels R.;
 RT "Isoform diversity of the inositol trisphosphate receptor in cell
 types of mouse origin.";
 RL Biochem. J. 322:575-583(1997).
 RN [3]
 RP SEQUENCE OF 818-1226 FROM N.A.
 RC STRAIN-C3H; TISSUE-Embryo;
 RX MEDLINE-94342363; PubMed-8063813;
 RA De Smedt H., Misiashen L., Parys J.B., Bootman M.D., Mertens L.,
 RA Van Den Bosch L., Casteels R.;
 RT "Determination of relative amounts of inositol trisphosphate receptor
 mRNA isoforms by ratio polymerase chain reaction.";
 RL J. Biol. Chem. 269:21691-21698(1994).
 CC -1- FUNCTION: RECEPTOR FOR INOSITOL 1,4,5-TRISPHOSPHATE, A SECOND
 CC MESSENGER THAT MEDIATES THE RELEASE OF INTRACELLULAR CALCIUM.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
 CC reticulum (By similarity).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM/TRIPR. ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: THE SHORT ISOFORM IS FOUND IN SKELETAL MUSCLE
 CC AND HEART.
 CC -1- DOMAIN: THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL
 CC EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-
 CC BINDING SITE IN THE N-TERMINAL AND MODULATORY SITES IN THE MIDDLE-
 CC PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.
 CC -1- PPM: PHOSPHORYLATED ON TYROSINE RESIDUES (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INSP3 RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

or send an email to license@isb-sib.ch.

CC
 CC EMBL: AB012393; BAA33960.1; -
 DR EMBL: 271173; CAA94861.1; -
 DR EMBL: 233908; CAA83957.1; -
 DR MGI: 99418; Tpr2.
 DR InterPro: IPR001682; Ca/Na.pore.
 DR InterPro: IPR000493; InsP3_receptor.
 DR InterPro: IPR000636; M-channel_nlg.
 DR InterPro: IPR003608; MIR.
 DR Pfam: PF00520; Ion_trans; 2.
 DR Pfam: PF02815; MIR; 1.
 DR PRINTS: PR00779; INSP3RECEPT.
 DR SMART: SM00472; MIR; 3.
 DR Receptor; Transmembrane; Glycoprotein; Phosphorylation;
 KW Endoplasmic reticulum; Ionic channel; Ion transport; Calcium channel;
 KW Alternative splicing.
 FT NON_CONS 272 273
 FT DOMAIN 1 807
 FT TRANSMEM 808 828
 FT DOMAIN 829 840
 FT TRANSMEM 841 861
 FT DOMAIN 862 887
 FT TRANSMEM 888 908
 FT DOMAIN 909 931
 FT TRANSMEM 932 952
 FT DOMAIN 953 974
 FT TRANSMEM 975 995
 FT DOMAIN 996 1101
 FT TRANSMEM 1102 1122
 FT DOMAIN 1123 1281
 FT MOD_RES 1187 1187
 FT VARSPLIT 175 175
 FT VARSPLIT 176 1281
 SO SEQUENCE 1281 AA; 145073 MW; FA87899B659A5598 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 1281;
 Best Local Similarity 85.7%; Pred. No. 2e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ELEDKON 7
 DB 97 ELEDKON 103

RESULT 14
 HMW2_MYCPN
 ID HMW2_MYCPN STANDARD; PRT: 1818 AA.
 AC P75471;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome c high molecular weight protein 2 (Cytochrome accessory
 DE protein 2).
 DE HMW2 OR MPN310 OR MP526.
 GN Mycoplasma pneumoniae.
 OS Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OC NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 29342 / M129;
 RX MEDLINE-97105885; PubMed-8948633;
 RA Himmelfreisch R., Hilbert H., Plagens H., Parkl E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae.";
 RN Nucleic Acids Res. 24:4420-4449(1996).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 29342 / M129;
 RX MEDLINE-97252497; PubMed-9098066;
 RA Krause D.C., Proft T., Hedreya C.T., Hilbert H., Plagens H.,
 RA Herrmann R.;

RT "Transposon mutagenesis reinforces the correlation between Mycoplasma pneumoniae cytoskeletal protein HMW2 and cytoadherence.";
 RL J. Bacteriol. 179:2668-2677(1997).
 CC -1- FUNCTION: COMPONENT OF THE CYTOSKELETON-LIKE STRUCTURE WHICH
 CC STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS
 CC CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMW
 CC PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYTADHERESIN PROTEINS
 CC IN THE MYCOPLASMA MEMBRANE AT THE ATTACHMENT ORGANELLE (BY
 CC SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: AE000051; AAB96174.1; -
 CC DR EMBL: U59896; AAB52527.1; -
 CC DR PhosSite; P75471; -
 CC DR Cytoadherence; Structural protein; Coiled coil; Complete proteome.
 CC FT DOMAIN 31 880 COILED COIL (POTENTIAL).
 CC FT DOMAIN 919 1607 COILED COIL (POTENTIAL).
 CC FT DOMAIN 1644 1755 COILED COIL (POTENTIAL).
 CC FT DOMAIN 1786 1817 COILED COIL (POTENTIAL).
 CC FT SEQUENCE 1818 AA; 215622 MW; 66DF4B08F0FCB6C0 CRC64;
 SQ
 Query Match 83.3%; Score 30; DB 1; Length 1818;
 Best Local Similarity 71.4%; Pred. No. 2.9e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 ELEKRON 7
 Db 534 ELEKRON 540

CC reticulum.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM/TYPE; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: THE SHORT ISOFORM IS FOUND IN SKELETAL MUSCLE
 CC AND HEART.
 CC -1- DOMAIN: THE RECEPTOR CONTAINS A CALCIUM CHANNEL IN ITS C-TERMINAL
 CC EXTREMITY. ITS LARGE N-TERMINAL CYTOPLASMIC REGION HAS THE LIGAND-
 CC BINDING SITE IN THE N-TERMINUS AND MODULATORY SITES IN THE MIDDLE
 CC PORTION IMMEDIATELY UPSTREAM OF THE CHANNEL REGION.
 CC -1- PHOSPHORYLATED ON TYROSINE RESIDUES (BY SIMILARITY).
 CC -1- MISCELLANEOUS: CALCIUM APPEARS TO INHIBIT LIGAND BINDING TO THE
 CC RECEPTOR. MOST PROBABLY BY INTERACTING WITH A DISTINCT
 CC CALCIUM-BINDING PROTEIN WHICH THEN INHIBITS THE RECEPTOR.
 CC -1- SIMILARITY: BELONGS TO THE INSP3 RECEPTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL: D26350; BAA0384.1; -
 CC DR EMBL: AB012610; BAA33961.1; -
 CC DR Genew: HGNC:6181; ITPR2.
 CC DR ITPR2; 600144; -
 CC DR InterPro: IPR000699; Ca-rel_channel.
 CC DR InterPro: IPR001682; Ca/Na_pore.
 CC DR InterPro: IPR000493; Insp3_receptor.
 CC DR InterPro: IPR000636; M+channel_nlg.
 CC DR InterPro: IPR003608; MIR.
 CC DR Pfam: PF00520; Ion_trans_1.
 CC DR Pfam: PF01365; RYR-ITPR. 2.
 CC DR Pfam: PF02815; MIR. 4.
 CC DR PRINTS; PR00779; INSP3RECEPTR.
 CC DR SMART; SM00472; MIR. 4.
 CC DR Receptor; Transmembrane; Glycoprotein; Phosphorylation;
 CC Endoplasmic reticulum; Ionic channel; Ion transport; Calcium channel;
 CC Alternative splicing
 CC FT DOMAIN 1 2227
 CC FT TRANSMEM 2228 2248 CYTOPLASMIC (POTENTIAL).
 CC FT DOMAIN 2249 2260 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 2261 2281 POTENTIAL.
 CC FT TRANSMEM 2282 2307 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 2308 2328 POTENTIAL.
 CC FT DOMAIN 2329 2351 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 2352 2372 POTENTIAL.
 CC FT DOMAIN 2373 2394 CYTOPLASMIC (POTENTIAL).
 CC FT TRANSMEM 2395 2415 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 2416 2521 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 2522 2542 POTENTIAL.
 CC FT TRANSMEM 2543 2701 CYTOPLASMIC (POTENTIAL).
 CC FT MOD.RES 2607 2607 PHOSPHORYLATION (POTENTIAL).
 CC FT VARSPLIC 176 181 IYVGRK -> DASFWI (IN SHORT ISOFORM).
 CC FT VARSPLIC 182 2701 MISSING (IN SHORT ISOFORM).
 CC FT SEQUENCE 2701 AA; 308074 MW; EB5C7DDCD17F74A CRC64;
 SQ
 Query Match 83.3%; Score 30; DB 1; Length 2701;
 Best Local Similarity 85.7%; Pred. No. 4.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 ELEKRON 7
 Db 97 ELEKRON 103

Search completed: November 13, 2002, 13:16:52
 Job time : 6.80851 secs